PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/00	A2	(11) International Publication Number: WO 99/47557
		(43) International Publication Date: 23 September 1999 (23.09.99)

(21) International Application Number: PCT/US99/06051

(22) International Filing Date: 18 March 1999 (18.03.99)

(30) Priority Data:
60/078,634   18 March 1998 (18.03.98)   US

(71) Applicant: ONYX PHARMACEUTICALS, INC. [US/US]; 3031 Research Drive, Richmond, CA 94806 (US).

(72) Inventors: BOLLAG, Gideon; 172 Catalina Drive, Hercules, CA 94547 (US). HART, Matthew, J.; 1210 Masonic Avenue, Berkeley, CA 94706 (US). ROSCOE, William; 3099 California Street, San Francisco, CA 94115 (US). POLAKIS, Paul; 509 Barone Lane, Mill Valley, CA 94941 (US). STERNWEIS, Paul; 2103 Flat Creet Drive, Richardson, TX 75080 (US). KOZASA, Tohru; 7415 Centenary, Dallas, TX 75225 (US). JIANG, Xuejun; 2400 Waterview Parkway #932, Richardson, TX 75080 (US).

(74) Agent: GIOTTA, Gregory; Onyx Pharmaceuticals, Inc., 3031 Research Drive, Richmond, CA 94806 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: IDENTIFICATION OF FACTORS WHICH MEDIATE THE INTERACTION OF HETEROTRIMERIC G PROTEINS AND MONOMERIC G PROTEINS

(57) Abstract

Monomeric GTPase guanine nucleotide exchange factor (GEF) have been identified which also contain an RGS region analogous to those of GTPase activating proteins (GAP). One of these GEF proteins, a Rho GEF has been demonstrated to contain an RGS sequence that has GAP activity toward a α subunit of a heterotrimeric G prote in.

IDENTIFICATION OF FACTORS WHICH MEDIATE THE INTERACTION OF HETEROTRIMERIC G PROTEINS AND MONOMERIC G PROTEINS

BACKGROUND OF THE INVENTION

5   Signal transduction pathways linking extracellular factors to the activation of the Rho GTPase have been implicated in cell growth control and cytoskeletal rearrangements. Specifically, heterotrimeric G proteins have been shoαwn to mediate these pathways, although the mechanism of mediation has been unclear. The identification of factors which interact with both heterotrimeric G proteins and Rho GTPase would provide an important

10   tool for investigating and controlling various cell processes, including cell proliferative diseases.

SUMMARY OF THE INVENTION

The invention relates to a polypeptide, and corresponding nucleic acid. comprising an amino acid sequence of a novel RGS domain, obtainable, e.g., from a guanine nucleotide

15   exchange factor (GEF) protein, where the polypeptide preferably does not include a dbl homology (DH) domain or a pleckstrin homology (PH) domain. In a preferred embodiment, the polypeptide has GTPase activating activity and binding affinity for an a G protein subunit such as Gα.

The polypeptides and nucleic acids can be used as tools for research, therapeutics,

20   and diagnostics as discussed below.

The invention also relates to a method of identifying or assaying for a molecule, or mixture of molecules, that regulate the binding of an RGS domain of a GEF protein to a substrate, e.g., a G protein subunit such as Gα . In one embodiment, the method involves incubating, under effective conditions, a polypeptide having an RGS domain of a GEF

25   polypeptide, and optionally having GEF activity, with a Gα subunit, or a fragment thereof, in the presence and/or absence of a test molecule; and determining whether the presence of the test molecule regulates the binding between the polypeptide and the subunit, or fragment thereof. As discussed later, various RGS-GEF polypeptides binding substrates can be utilized.

30   In addition, the invention relates to a method of identifying or assaying for a molecule, or mixture of molecules, that regulates a stimulatory effect of a polypeptide

comprising an RGS domain of a GEF protein on a polypeptide having a GTPase activity. In a preferred embodiment, the method comprises incubating a Gα subunit and a GEF protein, under effective conditions, in the presence and absence of a test molecule and determining whether the presence of the test molecule regulates the stimulatory effect of the GEF protein

5　on Gα subunit GTPase activity.

The invention also relates to a method of identifying or assaying for a molecule that specifically regulates a stimulatory effect of a first polypeptide, such as an activated Gα subunit, or polypeptide having GTPase activity, on a nucleotide exchange factor activity of a second polypeptide. The second polypeptide preferably comprises a RGS-GEF domain

10　obtainable from a GEF, and more preferably is a guanine nucleotide exchange factor (GEF) for a monomeric G protein. In one embodiment of the method, a first assay is conducted by incubating an activated Gα subunit with a GEF protein and a monomeric G protein in the presence and absence of a test molecule; a second assay is conducted: by incubating a GEF protein and a monomeric G protein in the presence and absence of the test molecule, and a

15　determination is made as to whether the molecule has a different effect when the first assay is compared to the second assay.

The invention further relates to a method of identifying or assaying for a molecule, or mixture of molecules, that mimics the stimulatory effect of an activated Gα subunit on GEF mediated nucleotide exchange of a monomeric G protein. In one example, such a

20　method comprises identifying a test compound that exhibits a binding affinity for an RGS domain of GEF proteins, incubating a GEF protein and monomeric G protein in the presence or absence of the test compound, determining whether the test compound exhibits a stimulatory effect on GEF mediated nucleotide exchange of a monomeric G protein.

The invention further relates to a method of identifying or assaying for a molecule,

25　or mixture of molecules, that mimics the stimulatory effect of an RGS domain of GEF polypeptide on Gα subunit GTPase activity. In one example, such a method comprises identifying a test compound that exhibits a binding affinity for a Gα subunit and incubating a GTP loaded Gα subunit in the presence or absence of the test compound to determine whether the test compound exhibits a stimulatory effect on GEF mediated nucleotide

30　exchange of a monomeric G protein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1, Panel A depicts the alignment of the sequences from RGS proteins and the N-terminal region of p115 Rho GEF as performed by Clustal W with a secondary structure mask of RGS4 to assign penalties for gaps. The RGS homologous sequences of Lsc,

5   KIAA380, and DrhoGEF2 were further added to this alignment by Clustal W and manual adjustments. The (a) symbols above RGS4 indicate the α helices of the RGS domain of RGS4. Dark shaded boxes indicate conserved residues of the hydrophobic core of the RGS structure. Lightly shaded boxes show other conserved residues. Asterisks mark the residues of RGS4 which contact $G\alpha_{i1}$. Primary sequences used in the alignment are the following:

10   rat RGS4 (SwissProt accession number P49799), mouse RGS2 (O08849), human GAIP (P49795), rat RGS12 (O08774), rat RGS14 (O08773), human p115 (1654344), mouse Lsc (1389756), human KIAA380 (2224701) and Drosophila DrhoGEF2 (2760368).

Figure 1, Panel B depicts constructs of p115 Rho GEF that were employed in the studies described herein. Numbers indicate the residues of p115 in each construct. The

15   RGS, dbl(DH), and pleckstrin (PH) homology regions are indicated. GST=glutathione-S-transferase.

Figure 2, Panel A is a graph showing the hydrolysis of GTP bound to $G\alpha_{13}$ and $G\alpha_{12}$ at 15°C either with (●○) or without (■□) 10 nM p115 Rho GEF.

Figure 2, Panel B is a graph showing the hydrolysis at 4°C of GTP bound to $G\alpha_{13}$

20   (●) and $G\alpha_{12}$ (○) and in the presence of various concentrations of p115 Rho GEF. The initial rates of reaction were plotted as a function of the concentration of p115 Rho GEF.

Figure 3 is a graph showing the hydrolysis at 15°C of GTP bound to $G\alpha_{13}$ and $G\alpha_{12}$ with either full-length p115 Rho GEF (●), ΔNp115 (■), or RGS-p115 (▲), or without any p115 construct (▼).

25   **Figure 4** is a graph showing the hydrolysis of GTP bound to $G\alpha_{i1}$, $G\alpha_z$, $G\alpha_q$, and $G\alpha_s$ with 100 nM p115 Rho GEF (Δ), 100 nM RGS4 (□), or buffer control (○). Assays were performed at 4°C for $G\alpha_{i1}$ and $G\alpha_s$, at 15°C for $G\alpha_z$, and at 20°C for $G\alpha_q$.

Figure 5 is a graph showing the selective inhibition of p115 GAP activity by the AlF_4- activated forms of $G\alpha$ subunits. Panel A: P115 (400 nM) was incubated on ice for 15

30   minutes with various $G\alpha$ subunits (400 nM) in the presence of 30 μM $AlCl_3$, 10 mM NaF,

3

and 10 mM $MgSO_4$. The mixture was diluted 20-fold, mixed with 0.3 nM $G\alpha_{12}$(GTP) and

the hydrolysis of bound GTP was measured after incubation at 15°C for 2 minutes. Panel B:

P115 (400 nM) was incubated with various concentrations of $G\alpha_{12}$(GDP-AlF_4^-) (●) or

$G\alpha_{13}$(GDP-AlF_4^-) (■) as described for Panel A. The mixture was diluted 20-fold, mixed

5   with 1nM $G\alpha_{13}$(GTP) at 4°C and the hydrolysis of bound GTP was assessed over time. The

initial rate of GTPase of $G\alpha_{13}$ was plotted against the final concentration of α subunit GDP-

AlF_4^-. The filled triangle indicates the rate of GTPase of $G\alpha_{13}$ without p115.

   Figure 6, Panel A is an image of an immunoblot showing the detection of myc-

tagged p115 Rho GEF expression in COS cells using an anti-myc antibody.

10   **Figure 6**, Panel B is an image of an immunoblot showing the detection of a

coimmunoprecipitate of p115 Rho GEF and $G\alpha_{13}$ using an anti-myc antibody.

   Figure 6, Panel C is an image of an immunoblot showing the detection of the

coimmunoprecipitate of p115 Rho GEF and $G\alpha_{13}$ using an anti-$G\alpha_{13}$ antibody.

   Figure 6, Panel D is an image of an immunoblot showing the detection of p115 Rho

15   GEF and $G\alpha_{13}$ binding when purified $G\alpha_{13}$ is added to immunoprecipitated p115 Rho GEF

when using an anti $G\alpha_{13}$ antibody.

   Figure 7, Panel A is a graph showing the dissociation of bound GDP from 100 nM

RhoA after 10 minutes in the presence or absence of 100 nm $G\alpha_{13}$ or $G\alpha_{12}$ and in the

presence of various concentrations of p115 Rho GEF as indicated.

20   **Figure 7**, Panel B is a graph showing the dissociation of GDP from 100 nM RhoA

after 10 minutes in the presence of 25 nm p115 Rho GEF and the indicated concentrations of

$G\alpha_{13}$ or $G\alpha_{12}$. Unstimulated dissociation of GDP from RhoA is indicated by the lower

dashed line.

   Figure 7, Panel C is a graph showing the dissociation of GDP from 100 nM RhoA

25   after 10 minutes of incubation with p115 Rho GEF and $G\alpha_{13}$ that had been treated with

AMF, GTPγS or GDPβS as indicated.

   Figure 7, Panel D is a graph showing the dissociation of of GDP from 100 nM

RhoA after 10 minutes of incubation with p115 Rho GEF (25 nM) and various Gα subunits

(100 nM) as indicated.

Figure 8, Panel A is a graph showing the association of 1 nM [^{32}P]GTP to 100 nM RhoA in the presence of the indicated concentrations of truncated for full-length p115 Rho GEF as measured by filtration after 30 minutes at 30°C.

Figure 8, Panel B is a graph showing the dissociation of [^{3}H]-GDP from 100 nM

5   RhoA after incubation for 10 minutes in the presence or absence of 25 nM p115 Rho GEF, 20 nM Gα_{13}, and 300 nM GST-RGSp115 as indicated.

Figure 8, Panel C is a graph showing the dissociation of [^{3}H]-GDP from 100 nM RhoA after incubation for 10 minutes in the presence 25 nM p115 Rho GEF and in the presence or absence of 25 nM Gα_{13} and the indicated concentrations of Gα_{12}.

10   **Figure 9**, Panel A is an image of an immunoblot showing the detection of myc-tagged KIAA380 (designated FL147) expression in COS cells using an anti-myc antibody.

Figure 9, Panel B is an image of an immunoblot showing the detection of a coimmunoprecipitate of KIAA380 (designated FL147) and Gα_{12} using an anti-Gα_{12} antibody.

15   **Figure 10** is the a listing of the amino acid sequence for p115 Rho GEF.. The RGS domain is shown by amino acids 45-170.

Figure 11 is a listing of the nucleic acid sequence for p115 Rho GEF. The RGS domain is encoded by nucleotides 187-564.

Figure 12 is a listing of the amino acid sequence for KIAA380. The RGS domain is

20   shown by amino acids 310-432.

Figure 13 is a listing of the nucleic acid sequence for KIAA380. The RGS domain is encoded by nucleotides 1673-2041.

Figure 14 is a listing of the amino acid sequence for Lsc. The RGS domain is shown by amino acids 43-168.

25   **Figure 15** is a listing of the nucleic acid sequence for Lsc. The RGS domain is encoded by nucleotides 218-595.

Figure 16 is a listing of the amino acid sequence for DRhoGEF2. The RGS domain is shown by amino acids 924-1053

Figure 17 is a listing of the nucleic acid sequence for DRhoGEF2. The RGS domain

30   is encoded by nucleotides 3185-3574.

5

Figure 18 is a homology alignment of the RGS region of several proteins, including GEF proteins with RGS domains (e.g. p115 Rho GEF, Lsc, KIAA380, DrhoGEF). The alignment was performed using the Clustal method with a PAM250 residue weight table.

DETAILED DESCRIPTION OF THE INVENTION

5   G proteins transduce signals from a large number of cell surface heptahelical receptors to various intracellular effectors. Each heterotrimeric G protein is composed of a guanine nucleotide-binding α subunit and a high-affinity dimer of β and γ subunits. $G\alpha$ subunits are commonly classified into four subfamilies (G_s, G_i, G_q, and G_{12}) based on their amino acid sequence homology and function (A.G. Gilman, *Annu. Rev. Biochem*, **56**, 615

10   (1987); Y. Kaziro et al., *Annu. Rev. Biochem.*, **60**, 349 (1991); Hepler and Gilman, *Trends Biochem. Sci.*, **17**, 383, (1992)). The G_{12} subfamily, consists of two identified members to date, G_{12} and G_{13}.

In accordance with the present invention, the identification of proteins having activity as both a GTPase activating protein (GAP) for the α subunit of a heterotrimeric G

15   protein and activity as a guanine nucleotide exchange factor (GEF) activity for monomeric G proteins have been described. Also in accordance with the invention, the first identification of a protein having GAP activity for the G_{12} subfamily of G proteins has been described. Also in accordance with the invention, the ability of an α subunit of a heterotrimeric G protein to stimulate GEF mediated guanine nucleotide exchange activity of a monomeric G

20   protein has been described. GAP and GEF activity, and methods of screening thereof, are described in Berman et al., 1996, *Cell* **86**:445 and Hart et al., 1996, *J. Biol. Chem.*, **271**:25452.

According to the present invention, the GAP activity of GEF proteins has been correlated with a novel RGS domain obtainable from a GEF protein. The present invention

25   relates to all aspects of such an RGS domain, including all aspects of a Rho GEF such as p115 Rho-GEF. (U.S. Patent Application No. 08/943,768, herein incorporated by reference).

A GEF protein modulates cell signaling pathways, both in *in vitro* and *in vivo*, by modulating the guanine nucleotide exchange activity of a GTPase. According to the present invention, a GEF protein which also modulates the GTPase activity of a heterotrimeric $G\alpha$

30   subunit is described. By way of illustration, p115 Rho-GEF, which modulates the guanine

nucleotide exchange activity of a Rho GTPase, as well as the GTPase activity of the $G\alpha_{12}$ family of heterotrimeric G protein subunits is described.

The present invention particularly relates to polypeptides comprising a RGS domain of a GEF polypeptide, or fragments thereof, and corresponding nucleic acids.

5   The invention also relates to methods of using such polypeptides, nucleic acids, or derivatives thereof, e.g., in therapeutics, diagnostics, and as research tools. Other aspects of the present invention relate to antibodies and other ligands which recognize the RGS domain of GEF polypeptides or nucleic acids, methods for identifying or assaying modulators of the GEF activities and/or the GAP activities of a protein containing a RGS domain, and

10   methods of treating pathological conditions associated with or related to the RGS domain, e.g., a GEF mediated interaction of a $G\alpha$ subunit and a Rho GTPase.

As used herein, an "RGS-GEF polypeptide" means, e.g., a polypeptide containing an RGS domain derived from a GEF protein, such as p115 Rho-GEF, Lsc, KIAA0380, or DRhoGEF2, and, which has one or more of the following activities: a specific binding

15   affinity for a polypeptide substrate, e.g., a G protein subunit, preferably an α subunit, such as G_{12} or G_{13}; a GTPase activating activity (GAP), such as a GAP activity for a G protein α subunit; or, an immunogenic activity. An RGS-GEF polypeptide preferably does not contain a (dbl homology) DH or a (pleckstrin homology) PH domain. DH and PH domains are disclosed in Cerione and Zheng, 1996, *Curr. Opin. In Cell Biol.*, 8:216. For example, the

20   amino acid sequence of p115 Rho GEF (Fig.10) contains a novel RGS domain at amino acids 45-170, the DH domain at amino acids 420-637, and the PH domain at amino acids 646-672. By "derived," it is meant that the amino acid sequence is obtainable from a naturally-occurring GEF (such as p115, Lsc, KIAA380, and DrrhoGEF2) or a non-naturally-occurring "mutated" sequence which is based upon a naturally-occurring GEF sequence

25   (i.e., different amino acid residues have been substituted for the amino acid residues which occur in the naturally-occurring sequence at a particular position). The polypeptide can be "isolated," i.e., the material is in a form in which it is not found in its original environment, e.g., more concentrated, more purified, or separated from other components,etc. A preferred RGS polypeptide possesses both a GAP and GEF activity, e.g., a mutated p115 Rho-GEF.

30   See below.

7

An RGS-GEF nucleic acid codes for an RGS-GEF polypeptide. The nucleic acid refers to both sense and anti-sense nucleic acids.

By the term "specific binding affinity," it is meant, e.g., that the RGS-GEF polypeptide has a binding preference for the activated state or transition state of a G protein

5   subunit as compared to a GDP-bound state or the nucleotide depleted state. By "GEF activity," it is meant, e.g., that the polypeptide stimulates or catalyzes the dissociation of GDP from a monomeric G-protein, such as Rho, and subsequent binding of GTP. Monomeric G-proteins include but are not limited to G-proteins in the Ras, Rho/Rac, Sar, Rab, Arf, and Ran families. Of particular interest are the RGS domains of the following

10   GEF proteins: human p115 (1654344) (Fig. 10, RGS domain at amino acids 45-170), mouse Lsc (1389756) (Fig. 14, RGS domain at amino acids 43-168), KIAA380 (2224701) (Fig. 12, RGS domain at amino acids 310-432) and Drosophila DrhoGEF2 (2760368) (Fig. 16, RGS domain at amino acids 924-1053).

Another aspect of the invention relates to novel consensus sequences for RGS

15   domain(s) of a GEF protein, herein referred to as a "sub-RGS consensus sequence." An "RGS domain," as used herein, refers to the amino acid sequence of protein which is able to bind to or physically interact with a G protein and, optionally, stimulates GTPase activity of that protein. A "sub-RGS consensus sequence," as used herein, refers to a consensus sequence which can be used to identify a specific subset of proteins which contain an RGS

20   domain. For example, a homology alignment of the RGS domain from several proteins as shown and described in Fig. 18 and the corresponding legend, shows that several sub-RGS consensus sequences may be defined by the gap of 13 to 14 amino acids that is apparent in the RGS domains of GEF proteins. One of these consensus sequences, herein designated as "RGS-GEF consensus 1," is herein defined to be a consensus sequence of AA_1-AA_2-AA_3-

25   AA_4-AA_5-AA_6-AA_7-AA_8-(gap of 13 amino acids)-AA_{22} -AA_{23}-AA_{24}-AA_{25}-AA_{26}, wherein:

AA_1 is L;

AA_2 is E or V;

AA_3 is K or P;

AA_4 is T, N, or R;

30   AA_5 is A;

AA_6 is V or P

AA$_7$ is L;

AA$_8$ is either S or a gap of one amino acid, contiguous with the gap of 13 amino acids;

AA$_{22}$ is either R or W;

AA$_{23}$ is either V or Y;

5   AA$_{24}$ is either P,K, or R

AA$_{25}$ is either V, I, or Q;

AA$_{26}$.is either P or D.

A second consensus sequence, herein designated as "RGS-GEF consensus 2," is

herein defined to be a consensus sequence of AA$_1$-AA$_2$-AA$_3$-AA$_4$-(gap of 13 amino acids)-

10   AA$_{18}$ -AA$_{19}$, wherein:

AA$_1$ is A;

AA$_2$ is V or P;

AA$_3$ is L;

AA$_4$ is either S or a gap of one amino acid, contiguous with the gap of 13 amino acids;

15   AA$_{18}$ is either R or W;

AA$_{19}$ is either V or Y.

Other proteins, including other GEF proteins can be aligned with the RGS domain of RGS

proteins as shown in Figure 18, and using methods described herein, to determine if they

contain a sub-RGS consensus sequence, such as RGS-GEF consensus 1 or RGS-GEF

20   consensus 2, as defined above.

In examining Figure 18 it is also apparent that a nucleotide sequence uniques to RGS

proteins that are not GEF proteins is shown by the nucleotide sequences which encode the

amino acids that correspond to the 13-14 amino acid gap in RGS-GEF proteins. These

nucleotide sequences could be used as probes to identify particular types of RGS proteins.

25   RGS-GEF polypeptides are preferably biologically-active. By biologically-active, it

is meant that a polypeptide fragment possesses an activity in a living system or with

component(s) of a living system. Biological-activities include, but are not limited to a

specific binding affinity for a G protein α subunit, as defined above, and GAP activity

toward a G protein α subunit. As described in the examples, such polypeptides can be

30   prepared routinely, e.g., by recombinant means or by proteolytic cleavage of isolated

polypeptides, and then assayed for a desired activity.

9

A polypeptide of the present invention includes polypeptides which have less than 100% identity to the amino acid sequences of p115 Rho-GEF (Fig. 10), Lsc (Fig. 14), KIAA0380 (Fig. 12), or DRhoGEF2 (Fig. 16). For the purposes of the following discussion: Sequence identity means that the same nucleotide or amino acid which is found in the

5   sequences set forth in Fig. 10-17 is found at the corresponding position of the compared sequence(s). A polypeptide having less than 100% sequence identity to the amino acid sequences set forth in Figures 10, 12, 14, and 16 can be substituted in various ways, e.g., by a conservative amino acid. The sum of the identical and conservatively substituted residues divided by the total number of residues in the sequence is equal to the percent sequence

10   similarity. For purposes of calculating sequence identity and similarity, the compared sequences can be aligned and calculated according to any desired method, algorithm, computer program, etc., including, e.g., FASTA, BLASTA. A polypeptide having less than 100% amino acid sequence identity to the amino acid sequences of the GEF proteins shown in Figures 10, 12, 14, and 16 may comprise, for example, about 60, 65 percent sequence

15   similarity and more preferably about 67, 70, 78, 80, 90, 92, 96, 99, etc. percent sequence amino acid sequence similarity.

In particular, the present invention relates to polypeptides, and corresponding nucleic acids, of p115, Lsc, KIAA380, and DrhoGEF2 which are mutated in the RGS domain of a GEF protein and which possess one or more of the RGS-GEF polypeptide activities

20   mentioned above. By the term "mutated," it is meant herein that such sequences are not naturally-occurring. For example a mutated polypeptide as mentioned can have one or more naturally-occurring positions replaced by a conservative amino acid, e.g., (based on the size of the side chain and degree of polarization) small nonpolar: cysteine, proline, alanine, threonine; small polar: serine, glycine, aspartate, asparagine; large polar: glutamate,

25   glutamine, lysine, arginine; intermediate polarity: tyrosine, histidine, tryptophan; large nonpolar: phenylalanine, methionine, leucine, isoleucine, valine. Such conservative substitutions also include those described by Dayhoff in the Atlas of Protein Sequence and Structure 5 (1978), and by Argos in EMBO J., 8, 779-785 (1989). A polypeptide having an amino acid sequence as set forth in Figures 10, 12, 14, and 16 can be substituted at 1, 5, 10,

30   15, or 20 positions by conservative amino acids. The mutations can be introduced into the conserved consensus region or the other residues of the RGS domain of a GEF protein.

A mutation to an RGS-GEF polypeptide can be selected to have one or more of the activities mentioned above, e.g., a specific binding affinity for a G protein α subunit, a GAP activity toward a G protein α subunit, etc.. Assays for such activities can be conducted as described below or as disclosed in Cerione and Zheng, 1996, *Curr. Opin. In Cell Biol.*,

5   8:216.

An RGS-GEF polypeptide can be modified by introducing amino acid substitutions into the hydrophobic core of the RGS domain (See Fig. 1, Panel A). For example, a conservative amino acid substitution would not be expected to affect activity, whereas as non-conservative amino acid substitution, e.g., changing a hydrophobic residue to a

10   hydrophilic residue, would be expected to reduce or eliminate its activity. Hydrophobic residues are nonpolar amino acids such as phenylalanine, leucine, isoleucine, valine, alanine, methionine, tryptophan, and cysteine. Hydrophilic residues are polar amino acids such as lysine, arginine, histidine, glutamate, and aspartate.

Modifications to a RGS-GEF polypeptide of the present invention or corresponding

15   nucleotide sequence, e.g., mutations, can also be prepared based on homology searching from gene data banks, e.g., Genbank, EMBL. Sequence homology searching can be accomplished using various methods, including algorithms described in the BLAST family of computer programs, the Smith-Waterman algorithm, etc. For example, conserved amino acids can be identified between various sequences containing an RGS domain of various

20   GEF proteins. (See Fig. 18) A mutation(s) can then be introduced into such sequences by identifying and aligning amino acids conserved between the polypeptides and then modifying an amino acid in a conserved or non-conserved position. A mutated RGS-GEF sequence can comprise conserved or non-conserved amino acids, e.g., between corresponding regions of homologous nucleic acids. For example, a mutated sequence can

25   comprise conserved or non-conserved residues from any number of homologous sequences as mentioned-above and/or determined from an appropriate searching algorithm.

Corresponding mutations can be made in specific regions of an RGS-GEF nucleic acid. For example, mutations may be made wherein amino acids that particpate in the GTPase catalytic function or mutations may be made in amino acids that function as contact

30   points between the RGS-GEF sequence and the Gα subunit.

An RGS-GEF polypeptide or fragment thereof, or substituted RGS-GEF polypeptide or fragment thereof, may also comprise various modifications, wherein such modifications include glycosylation, covalent modifications (e.g., of an R-group of an amino acid), amino acid substitution, amino acid deletion, or amino acid addition. Modifications to the

5   polypeptide can be accomplished according to various methods, including recombinant, synthetic, chemical, etc.

Polypeptides of the present invention (e.g., RGS-GEF polypeptides, and fragments and mutations thereof) may be used in various ways, e.g., as immunogens for antibodies as described below, as biologically-active agents (e.g., having one or more of the activities

10   associated with an RGS-GEF polypeptide), as inhibitors of the activities of the corresponding full-length polypeptide. For example, upon binding of p115 Rho-GEF to the Gα subunit, a cascade of events is initiated in the cell, e.g., promoting cell proliferation and/or cytoskeletal rearrangements. The interaction between p115 Rho-GEF and the Gα subunit can be modulated by using a RGS-GEF polypeptide, or fragment thereof, to inhibit

15   the interaction between p115 Rho-GEF and the Gα subunit. Such a fragment can be useful for modulating pathological conditions associated with the Rho signaling pathway. A useful fragment may be identified routinely by testing the ability of overlapping fragments of the entire length of the RGS domain of a GEF protein to inhibit the binding of p115 Rho-GEF with the Gα subunit or to inhibit the GAP activity of the p115 Rho-GEF toward the Gα

20   subunit. The measurement of these activities is described below and in the examples. Peptides can be chemically-modified, etc.

A RGS-GEF polypeptide of the present invention can comprise one or more structural domains, functional domains, detectable domains, antigenic domains, and/or other polypeptides of interest, in an arrangement which does not occur in nature, i.e., not

25   naturally-occurring. A polypeptide comprising such features is a chimeric or fusion polypeptide. Such a chimeric polypeptide can be prepared according to various methods, including, chemical, synthetic, quasi-synthetic, and/or recombinant methods. A chimeric nucleic acid coding for a chimeric polypeptide can contain the various domains or desired polypeptides in a continuous or interrupted open reading frame, e.g., containing introns,

30   splice sites, enhancers, etc. The chimeric nucleic acid can be produced according to various methods. See, e.g., U.S. Pat. No. 5,439,819. A domain or desired polypeptide can possess

12

any desired property, including, a biological function such as catalytic, signaling, growth

promoting, cellular targeting, etc., a structural function such as hydrophobic, hydrophilic,

membrane-spanning, etc., receptor-ligand functions, and/or detectable functions, e.g.,

combined with enzyme, fluorescent polypeptide, green fluorescent protein GFP (Chalfie et

5   al., 1994, *Science*, 263:802; Cheng et al., 1996, *Nature Biotechnology*, 14:606; Levy et al.,

1996, *Nature Biotechnology*, 14:610, etc. In addition, a RGS-GEF nucleic acid, or a

fragment thereof, may be used as selectable marker when introduced into a host cell. For

example, a nucleic acid coding for an amino acid sequence according to the present

invention can be fused in frame to a desired coding sequence and act as a tag for

10   purification, selection, or marking purposes. The region of fusion encodes a cleavage site.

            A polypeptide according to the present invention can be produced in an expression

system, e.g., in vivo, in vitro, cell-free, recombinant, cell fusion, etc., according to the

present invention. Modifications to the polypeptide imparted by such system include,

glycosylation, amino acid substitution (e.g., by differing codon usage), polypeptide

15   processing such as digestion, cleavage, endopeptidase or exopeptidase activity, attachment

of chemical moieties, including lipids, phosphates, etc. For example, some cell lines can

remove the terminal methionine from an expressed polypeptide.

            A polypeptide according to the present invention can be recovered from natural

sources, transformed host cells (culture medium or cells) according to the usual methods,

20   including, ammonium sulfate or ethanol precipitation, acid extraction, anion or cation

exchange chromatography, phosphocellulose chromatography, hydrophobic interaction

chromatography, hydroxyapatite chromatography and lectin chromatography. It may be

useful to have low concentrations (approximately 0.1-5 mM) of calcium ion present during

purification (Price, et al., *J. Biol. Chem.*, 244:917 (1969)). High performance liquid

25   chromatography (HPLC) can be employed for final purification steps.

            A RGS-GEF nucleic acid of the present invention can comprise the complete coding

sequence for an RGS-GEF polypeptide, or fragments thereof. A nucleic acid according to

the present invention may also comprise a nucleotide sequence which is 100%

complementary, e.g., an anti-sense, to any RGS-GEF nucleotide sequence.

30   A nucleic acid according to the present invention can be obtained from a variety of

different sources. It may be obtained from DNA or RNA, such as polyadenylated mRNA,

13

e.g., isolated from tissues, cells, or whole organism. The nucleic acid may be obtained
directly from DNA or RNA, or from a cDNA library. The nucleic acid can be obtained from
a cell at a particular stage of development, having a desired genotype, phenotype (e.g., an
oncogenically transformed cell or a cancerous cell), etc. The nucleic acid may also be
5   chemically synthesized.

        A nucleic acid according to the present invention may include only coding sequence
for an RGS-GEF polypeptide; coding sequence for an RGS-GEF polypeptide and additional
functional coding sequences including, for example, leader sequences, secretory sequences,
tag sequences (e.g. targeting tags, enzymatic tags, fluorescent tags etc.). A nucleic acid
10   according to the present invention may also include coding sequence for an RGS-GEF
polypeptide and non-coding sequences, e.g., untranslated sequences at either a 5' or 3' end,
or dispersed in the coding sequence, e.g., introns.

        A nucleic acid according to the present invention may also comprise an expression
control sequence operably linked to a nucleic acid as described above. The phrase
15   "expression control sequence" means a nucleic acid sequence which regulates expression of
a polypeptide coded for by a nucleic acid to which it is operably linked. Expression can be
regulated at the level of the mRNA or polypeptide. Thus, the expression control sequence
includes mRNA-related elements and protein-related elements. Such elements include
promoters, enhancers (viral or cellular), ribosome binding sequences, transcriptional
20   terminators, etc. An expression control sequence is operably linked to a nucleotide coding
sequence when the expression control sequence is positioned in such a manner to effect or
achieve expression of the coding sequence. For example, when a promoter is operably
linked 5' to a coding sequence, expression of the coding sequence is driven by the promoter.
Expression control sequences can be heterologous or endogenous to the normal gene.

25   A nucleic acid in accordance with the present invention may be selected on the basis
of nucleic acid hybridization. The ability of two single-stranded nucleic acid preparations to
hybridize together is a measure of their nucleotide sequence complementarity, e.g., base-
pairing between nucleotides, such as A-T, G-C, etc. The invention thus also relates to
nucleic acids which hybridize to a nucleic acids comprising a nucleotide sequence as set
30   forth in Figures 11, 13, 15, and 17. The present invention includes both strands of nucleic
acid, e.g., a sense strand and an anti-sense strand.

According to the present invention, a nucleic acid or polypeptide can comprise one or more differences in the nucleotide or amino acid sequence set forth in Figures 10-17. Changes or modifications to the nucleotide and/or amino acid sequence can be accomplished by any method available, including directed or random mutagenesis.

5   A nucleic acid coding for an RGS-GEF polypeptide according to the invention may comprise nucleotides which occur in a naturally-occurring GEF gene e.g., naturally-occurring polymorphisms, normal or mutant alleles (nucleotide or amino acid), mutations which are discovered in a natural population of mammals, such as humans, monkeys, pigs, mice, rats, or rabbits. By the term naturally-occurring, it is meant that the nucleic acid is

10   obtained from a natural source, e.g., animal tissues and cells, body fluids, tissue culture cells, forensic samples. Naturally-occurring mutations include deletions, substitutions, or additions of nucleotide sequence. These genes can be detected and isolated by nucleic acid hybridization according to methods well known to one skilled in the art. It is recognized that, by analogy to other oncogenes, naturally-occurring variants of GEF proteins will

15   include variants with deletions, substitutions, and additions in the RGS domain of a GEF protein, which produce pathological conditions in the host cell and organism.

A nucleotide sequence coding for an RGS-GEF polypeptide of the invention may contain codons found in a naturally-occurring gene, transcript, or cDNA, for example, or it may contain degenerate codons coding for the same amino acid sequences.

20   In addition, a nucleic acid or polypeptide of the present invention may be obtained from any desired mammalian organism, but also non-mammalian organisms. Homologs from mammalian and non-mammalian organisms can be obtained according to various methods. For example, hybridization with an oligonucleotide (see below) selective for an RGS domain of a GEF, or a RGS-GEF, of the present invention can be employed to select

25   such homologs, e.g., as described in Sambrook et al., *Molecular Cloning*, 1989, Chapter 11. Such homologs may have varying amounts of nucleotide and amino acid sequence identity and similarity to previously identified RGS domain or RGS-GEF nucleotide or polypeptide sequence. Non-mammalian organisms include, e.g., vertebrates, invertebrates, zebra fish, chicken, *Drosophila*, yeasts (such as *Saccharomyces cerevisiae*), *C. elegans*, roundworms,

30   prokaryotes, plants, *Arabidopsis*, viruses, etc.

A nucleic acid according to the present invention may comprise, for example, DNA, RNA, synthetic nucleic acid, peptide nucleic acid, modified nucleotides, or mixtures thereof. A DNA can be double- or single-stranded. Nucleotides comprising a nucleic acid can be joined via various known linkages such as, for example, ester, sulfamate, sulfamide,

5   phosphorothioate, phosphoramidate, methylphosphonate, carbamate, etc., depending on the desired purpose. Linkages may be modified for purposes such as, for example, resistance to nucleases such as RNase H and improved *in vivo* stability. See, e.g., U.S. Pat. Nos. 5,378,825.

Various modifications can be made to the nucleic acids, such as attaching detectable

10   markers (avidin, biotin, radioactive elements), moieties which improve hybridization, detection, or stability. The nucleic acids can also be attached to solid supports, e.g., nitrocellulose, nylon, agarose, diazotized cellulose, latex solid microspheres, polyacrylamides, etc., according to a desired method. See, e.g., U.S. Pat. Nos. 5,470,967, 5,476,925, 5,478,893.

15   Another aspect of the present invention relates to oligonucleotides and nucleic acid probes. Such oligonucleotides or nucleic acid probes can be used, e.g., to detect, quantify, or isolate an RGS-GEF nucleic acid in a test sample. Detection can be desirable for a variety of different purposes, including research, diagnostic, and forensic. For diagnostic purposes, it may be desirable to identify the presence or quantity of a specific RGS-GEF

20   nucleic acid sequence in a sample obtained from tissues, cells, body fluids, etc. In a preferred method, the present invention relates to a method of detecting a target RGS-GEF nucleic acid in a test sample comprising contacting the test sample with an oligonucleotide under conditions effective to achieve hybridization between the target and oligonucleotide; and detecting hybridization. An oligonucleotide in accordance with the invention can also

25   be used in synthetic nucleic acid amplification such as PCR, e.g., Saiki et al., 1988, Science, 241:53; U.S. Pat. No. 4,683,20, or or differential display (See, e.g., Liang et al., *Nucl. Acid. Res.*, 21:3269-3275, 1993; USP 5,599,672; WO97/18454). Oligonucleotides can be identified routinely, e.g., to the DH, PH, and RGS-GEF domains to differentially display and/or amplify gene products containing such sequences.

30   Both sense and antisense nucleotide sequences are intended as part of the invention. A unique nucleic acid according to the present invention may be determined routinely. An

RGS-GEF nucleic acid may be used as a hybridization probe to identify the presence of RGS-GEF nucleotide sequence in a sample comprising a mixture of nucleic acids, e.g., on a Northern blot. Hybridization can be performed under stringent conditions to select nucleic acids having at least 95% identity (i.e., complementarity) to the probe, but less stringent

5   conditions can also be used. A unique RGS-GEF nucleotide sequence can also be fused in-frame, at either its 5' or 3' end, to various nucleotide sequences, including, for example, coding sequences for enzymes or expression control sequences, etc.

Hybridization can be performed under different conditions, depending on the desired selectivity, e.g., as described in Sambrook et al., *Molecular Cloning*, 1989. For example, to

10   specifically detect RGS-GEF sequences, an oligonucleotide can be hybridized to a target nucleic acid under conditions in which the oligonucleotide only hybridizes to the GEF sequence from which the RGS -GEF sequence was derived, e.g., where the oligonucleotide is 100% complementary to the target. Different conditions can be used if it is desired to select target nucleic acids which have less than 100% nucleotide complementarity, at least

15   about, e.g., 99%, 97%, 95%, 90%, 70%, 67%. Since a mutation in GEF genes can cause diseases or pathological conditions, e.g., cancer, benign tumors, an oligonucleotide according to the present invention can be used diagnostically. For example, a patient having symptoms of a cancer or other condition associated with the Rho signaling pathway (see below) can be diagnosed with the disease by using an oligonucleotide according to the

20   present invention, in polymerase chain reaction followed by DNA sequencing to identify whether the sequence is normal, in combination with other oncogene oligonucleotides, etc., e.g., p53, Rb, p21, Dbl, MTS1, Wt1, Bcl-1, Bcl-2, MDM2, etc.

Oligonucleotides according to the present invention can be of any desired size, preferably 14-16 oligonucleotides in length, or more. Such oligonucleotides can have non-

25   naturally-occurring nucleotides, e.g., inosine. In accordance with the present invention, the oligonucleotide can comprise a kit, where the kit includes a desired buffer (e.g., phosphate, tris, etc.), detection compositions, etc. The oligonucleotide can be labeled or unlabeled, with radioactive or non-radioactive labels as known in the art.

Anti-sense nucleic acid can also be prepared from a nucleic acid according to the

30   present, preferably an anti-sense RGS-GEF nucleotide sequence corresponding to an RGS-GEF nucleotide sequence of Figures 11, 13, 15, and 17. Anti-sense RGS-GEF nucleic acid

can be used in various ways, such as to regulate or modulate expression of GEF proteins containing RGS domains or to detect expression of RGS-GEF proteins, including by *in situ* hybridization. For the purposes of regulating or modulating expression, an anti-sense oligonucleotide may be operably linked to an expression control sequence.

5   The RGS-GEF nucleic acids according to the present invention can be labelled according to any desired method. The nucleic acid can be labeled using radioactive tracers such as ^{32}P, ^{35}S, ^{125}I, ^{3}H, or ^{14}C, to mention only the most commonly used tracers. The radioactive labeling can be carried out according to any method such as, for example, terminal labeling at the 3' or 5' end using a radiolabeled nucleotide, polynucleotide kinase

10   (with or without dephosphorylation with a phosphatase) or a ligase (depending on the end to be labeled). A non-radioactive labeling can also be used, combining a nucleic acid of the present invention with residues having immunological properties (antigens, haptens), a specific affinity for certain reagents (ligands), properties enabling detectable enzyme reactions to be completed (enzymes or coenzymes, enzyme substrates, or other substances

15   involved in an enzymatic reaction), or characteristic physical properties, such as fluorescence or the emission or absorption of light at a desired wavelength, etc.

          An RGS-GEF nucleic acid according to the present invention, including oligo-nucleotides, anti-sense nucleic acid, etc., can be used to detect expression of RGS-GEF nucleic acids in whole organs, tissues, cells, etc., by various techniques, including Northern

20   blot, PCR, *in situ* hybridization, etc. Such nucleic acids can be particularly useful to detect disturbed expression, e.g., cell-specific and/or subcellular alterations of RGS-GEF expression. The levels of RGS-GEF proteins can be determined alone or in combination with other genes products (oncogenes such as p53, Rb, Wt1, etc.), transcripts, etc.

          A nucleic acid according to the present invention can be expressed in a variety of

25   different systems, *in vitro* and *in vivo*, according to the desired purpose. For example, a nucleic acid can be inserted into an expression vector, introduced into a desired host, and cultured under conditions effective to achieve expression of a polypeptide coded for the nucleic acid. Effective conditions includes any culture conditions which are suitable for achieving production of the polypeptide by the host cell, including effective temperatures,

30   pH, medias, additives to the media in which the host cell is cultured (e.g., additives which amplify or induce expression such as butyrate, or methotrexate if the coding nucleic acid is

adjacent to a dhfr gene), cyclohexamide, cell densities, culture dishes, etc. A nucleic acid
can be introduced into the cell by any effective method including, e.g., calcium phosphate
precipitation, electroporation, injection, DEAE-Dextran mediated transfection, fusion with
liposomes, and viral transfection. A cell into which a nucleic acid of the present invention

5   has been introduced is a transformed host cell. The nucleic acid can be extrachromosomal
or integrated into a chromosome(s) of the host cell. It can be stable or transient. An
expression vector is selected for its compatibility with the host cell. Host cells include,
mammalian cells, e.g., COS-7, CHO, HeLa, LTK, NIH 3T3, yeast, insect cells, such as Sf9
(*S. frugipeda*) and *Drosophila*, bacteria, such as *E. coli, Streptococcus sp., Bacillus sp.*,

10   yeast, fungal cells, plants, embryonic stem cells (e.g., mammalian, such as mouse or
human), cancer or tumor cells Sf9 expression can be accomplished in analogy to Graziani et
al., *Oncogene*, 7:229-235, 1992. Expression control sequences are similarly selected for
host compatibility and a desired purpose, e.g., high copy number, high amounts, induction,
amplification, controlled expression. Other sequences which can be employed include

15   enhancers such as from SV40, CMV, inducible promoters, cell-type specific elements, or
sequences which allow selective or specific cell expression.

A labelled polypeptide can be used, e.g., in binding assays, such as to identify
substances that bind or attach to p115 Rho-GEF, to track the movement of p115 Rho-GEF
in a cell, in an *in vitro, in vivo*, or *in situ* system, etc.

20   A nucleic acid or polypeptide of the present invention can also be substantially
purified. By substantially purified, it is meant that nucleic acid or polypeptide is separated
and is essentially free from other nucleic acids or polypeptides, i.e., the nucleic acid or
polypeptide is the primary and active constituent.

Another aspect of the present invention relates to the regulation of biological

25   pathways in which a RGS-GEF polypeptide is involved, particularly pathological conditions,
e.g., cell proliferation (e.g., cancer), growth control, morphogenesis, stress fiber formation,
and integrin-mediated interactions, such as embryonic development, tumor cell growth and
metastasis, programmed cell death, hemostasis, leucocyte homing and activation, bone
resorption, clot retraction, and the response of cells to mechanical stress. See, e.g., Clark

30   and Brugge, Science, 268:233- 239, 1995; Bussey, Science, 272:225-226, 1996. Thus, the
invention relates to all aspects of a method of modulating an activity of a RGS-GEF

19

polypeptide comprising, administering an effective amount of an RGS-GEF polypeptide or a biologically-active fragment thereof, an effective amount of a compound which modulates the activity of a RGS-GEF polypeptide, or an effective amount of a nucleic acid which codes for a RGS-GEF polypeptide or a biologically-active fragment thereof. The activity of the

5   RGS-GEF which is modulated may include binding to a Gα subunit or GAP activity toward a Gα subunit. The activity can be modulated by increasing, reducing, antagonizing, or promoting expression or activity of the RGS-GEF.

The present invention also relates to antibodies which specifically recognize a RGS-GEF polypeptide. Antibodies, e.g., polyclonal, monoclonal, recombinant, chimeric, can be

10   prepared according to any desired method. For example, for the production of monoclonal antibodies, an RGS-GEF polypeptide according to Figures 10, 12, 14, or 16 can be administered to mice, goats, or rabbit subcutaneously and/or intraperitoneally, with or without adjuvant, in an amount effective to elicit an immune response. The antibodies can also be single chain or FAb. The antibodies can be IgG, subtypes, IgG2a, IgG1, etc.

15   An antibody specific for RGS-GEF means that the antibody recognizes a defined sequence of amino acids within or including the amino acid sequence of the RGS domain of a GEF polypeptide. Thus, a specific antibody will bind with higher affinity to an amino acid sequence, i.e., an epitope, found in the RGS domain of a GEF polypeptdie than to a different epitope(s), e.g., as detected and/or measured by an immunoblot assay. Thus, an antibody

20   which is specific for an epitope within or including the RGS domain of p115 Rho-GEF is useful to detect the presence of the epitope in a sample, e.g., a sample of tissue containing p115 Rho-GEF gene product, distinguishing it from samples in which the epitope is absent.

Additionally, in accordance with the present invention, ligands which bind to an RGS domain of a GEF polypeptide can also be prepared, e.g., using synthetic peptide

25   libraries or aptamers (e.g., Pitrung et al., U.S. Pat. No. 5,143,854; Geysen et al., 1987, J. Immunol. Methods, 102:259-274; Scott et al., 1990, Science, 249:386; Blackwell et al., 1990, Science, 250:1104; Tuerk et al., 1990, Science, 249: 505.

Antibodies and other ligands which bind the RGS domain of a GEF polypeptide, and specifically antibodies and other ligands which bind the RGS domain of p115 Rho GEF, can

30   be used in various ways. These include, but are not limited to, uses therapeutic, diagnostic, and commercial research tools, e.g, to quantitate the levels of p115 Rho-GEF polypeptide in

animals, tissues, cells, etc., to identify the cellular localization and/or distribution of p115
Rho-GEF, to purify p115 Rho-GEF or a polypeptide comprising a part of p115 Rho-GEF, to
modulate the function of p115 Rho-GEF, etc. Antibodies can be used in Western blots,
ELIZA, immunoprecipitation, RIA, etc. The present invention relates to such assays,
5   compositions and kits for performing them, etc.

An antibody according to the present invention can be used to detect polypeptides or
fragments containing an RGS domain of a GEF polypeptide in various samples, including
tissue, cells, body fluid, blood, urine, cerebrospinal fluid. A method of the present invention
comprises contacting a ligand which binds to an RGS-GEF polypeptide of Figure 10, 12, 14,
10   or 16 under conditions effective, as known in the art, to achieve binding, detecting specific
binding between the ligand and peptide. By specific binding, it is meant that the ligand
attaches to a defined sequence of amino acids, e.g., within or including the amino acid
sequence of the RGS domain as shown in Figures 10, 12, 14, and 16, or derivatives thereof.
The antibodies or derivatives thereof can also be used to inhibit expression of GEF proteins
15   containing an RGS domain. The levels of a GEF polypeptide containing an RGS domain
may be determined alone or in combination with other gene products. In particular, the
amount (e.g., its expression level) of the GEF polypeptide containing an RGS domain can be
compared (e.g., as a ratio) to the amounts of other polypeptides in the same or different
sample, e.g., p21, p53, Rb, WT1, etc.

20   A ligand for the RGS domain of GEF polypeptides can be used in combination with
other antibodies, e.g., antibodies that recognize oncological markers of cancer, including,
Rb, p53, c-erbB-2, oncogene products, etc. In general, reagents which are specific for the
RGS domain of GEF polypeptides can be used in diagnostic and/or forensic studies
according to any desired method, e.g., as U.S. Pat. Nos. 5,397,712; 5,434,050; 5,429,947.

25   The present invention also relates to a transgenic animal, e.g., a non-human-
mammal, such as a mouse, comprising an RGS-GEF polypeptide. Transgenic animals can
be prepared according to known methods, including, e.g., by pronuclear injection of
recombinant genes into pronuclei of 1-cell embryos, incorporating an artificial yeast
chromosome into embryonic stem cells, gene targeting methods, embryonic stem cell
30   methodology. See, e.g., U.S. Patent Nos. 4,736,866; 4,873,191; 4,873,316; 5,082,779;
5,304,489; 5,174,986; 5,175,384; 5,175,385; 5,221,778; Gordon et al., *Proc. Natl. Acad.*

21

Sci., 77:7380-7384 (1980); Palmiter et al., *Cell*, 41:343-345 (1985); Palmiter et al., *Ann. Rev. Genet.*, 20:465-499 (1986); Askew et al., *Mol. Cell. Bio.*, 13:4115-4124, 1993; Games et al. *Nature*, 373:523-527, 1995; Valancius and Smithies, *Mol. Cell. Bio.*, 11:1402-1408, 1991; Stacey et al., *Mol. Cell. Bio.*, 14:1009-1016, 1994; Hasty et al., *Nature*, 350:243-246,

5   1995; Rubinstein et al., *Nucl. Acid Res.*, 21:2613-2617,1993. A nucleic acid according to the present invention can be introduced into any non-human mammal, including a mouse (Hogan et al., 1986, in *Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), pig (Hammer et al., *Nature*, 315:343-345, 1985), sheep (Hammer et al., *Nature*, 315:343-345, 1985), cattle, rat, or

10   primate. See also, e.g., Church, 1987, *Trends in Biotech.* 5:13-19; Clark et al., 1987, *Trends in Biotech.* 5:20-24, and DePamphilis et al., 1988, *BioTechniques*, 6:662-680. Additionally, custom transgenic rat and mouse production is commercially available. These transgenic animals are useful, for example, as a cancer model or as a model to evaluate the effects of overexpression of the RGS-GEF polypeptide.

15   Generally, the nucleic acids, polypeptides, antibodies, etc. of the present invention can be prepared and used as described in, U.S. Pat. Nos. 5,501,969, 5,506,133, 5,441,870; WO 90/00607; WO 91/15582;

Other aspects of this invention relate to methods to assay for, or identify, molecules that modulate the following interactions and effects: the interaction between an RGS domain

20   of a GEFand its cognate binding substrate; the interaction between an RGS domain of a GEF and a Gα subunit; the effect of G protein subunit stimulation on a guanine nucleotide exchange activity of a GEF protein containing an RGS domain; the effect of a GEF protein having an RGS domain as a GTPase activating protein for a G protein subunit.

Activity can be modulated in various ways, e.g., enhancing, activating, stimulating,

25   suppressing, preventing, inhibiting, etc. A modulatory molecule can be an agonist, antagonist, or have partial activities thereof. Modulating molecules can be any type of molecule, including but not limited to small molecules, proteins, peptides, antibodies, nucleic acids, etc. In general, a compound having an *in vitro* activity will be useful *in vivo* to modulate a biological pathway associated with a GEF protein containing an RGS domain,

30   e.g., to treat a pathological condition associated with the biological and cellular activities

mentioned above. The modulatory molecules can comprise a mixture of the same or different molecules.

A binding substrate for the RGS domain of a GEF protein can be any material to which the RGS domain binds specifically, including members of the $G\alpha 12$ family. See,

5   e.g., Strathman and Simon, *Proc. Natl. Acad. Sci.*, 88:5582, 1991. For example, a method of identifying or assaying for a molecule that modulates or regulates the binding of a G protein α subunit to a GEF protein containing an RGS domain, such as p115 Rho-GEF, can be conducted in accordance with this invention. In one embodiment, a GTP bound α subunit, or derivative thereof, is incubated with a GEF protein, or fragment thereof,

10 containing the RGS domain, in the presence and absence of a test molecule to determine whether the presence of the test compound modulates the binding between the GEF protein and the G protein α subunit. The incubation is accomplished under effective conditions, i.e., conditions under which binding or attachment occurs. Binding can be detected in one or more ways. For example, the GEF protein or the binding substrate is labeled detectably;

15 the labelled bound component is separated from the labelled free component; and the amount of bound-detectably labeled GEF protein or binding substrate determined. The detectable label can be of any desired composition, e.g., radioactive, fluorescent, etc. Such an assay can be performed in either solid or liquid phase.

In one aspect of the invention, it is desirable to identify molecules that regulate the

20 binding of the $G\alpha_{12}$ family of subunits, eg. $G\alpha_{12}$ and $G\alpha_{13}$, with a GEF, e.g., p115 Rho GEF, Lsc, KIAA380, or DrhoGEF2. The assay can be conducted using a complete GEF protein, or any subfragments thereof which contain the RGS domain, or biologically active subfragments of the RGS domain. The assay is typically conducted with stable analogs of the GTP bound state of the $G\alpha$ subunit, including α subunits bound to either GDP-AIF$_4^-$ or

25 GTPγS. For example, a binding assay may be conducted by the procedure described in Example 5 below wherein a COS cell is transfected with a nucleic acid construct for a myc-tagged polypeptide, such as p115 Rho GEF, or a fragment thereof and complexes of the polypetide and a $G\alpha$ subunit are detected by precipitation of any bound complex with a first antibody to one of components and detection of the amount of a second bound component

30 with a second antibody. Binding assays could also be performed using techniques that are

23

well known in the art such as by binding one of the components to a column and then determining the amount of a second labelled compoent that binds to the column. Relevant assay methods are also disclosed, for example, in Berman et al, 1996, *J. Biol. Chem.* 271:27209.

5   A method of isolating or assaying for a molecule that modules or regulates the stimulatory effect of a RGS-GEF polypeptide on GTPase activity, such as a GTPase activity of a Gα subunit, can also be conducted in accordance with the invention. For example, a Gα subunit is incubated under effective conditions with an RGS-GEF polypeptide having GTPase stimulatory effect in the presence and absence of a test inhibitor to determine

10   whether the presence of the test inhibitor modulates its stimulatory effect. The assay can conducted using a complete RGS-GEF polypeptide, a GEF protein, or any subfragments thereof which contain the RGS domain, or biologically active subfragments of the RGS domain. An RGS-GEF polypeptide can be p115 Rho GEF, Lsc, KIAA380, DrhoGEF2, or biologically-active fragments thereof. For example, an assay can be conducted using a p115

15   Rho GEF in conjuction with an α12 or an α13 subunit, as described in the examples discussed herein, as well as using other variations or assay methods which are well known in the art. For example, the assay may be conducted in accordance with Example 5 below, in which Gα subunits were loaded with $[\gamma\text{-}^{32}P]GTP$ and the amount of hydrolysis under various conditions, including the presence of an RGS-GEF polypeptide, was determined by

20   measuring the amount of ^{32}Pi in the supernatant after centrifugation of the assay mixture. Relevant assay methods are also disclosed, for example, in Berman et al, 1996, *J. Biol. Chem.* 271:27209.

         A method of identifying or assaying for a moiecule that modulates the stimulatory effect of an activated Gα subunit on a RGS-GEF polypeptide having GEF mediated

25   nucleotide exchange for a monomeric G protein can also be conducted in accordance with this invention. For instance, a first assay can be conducted by incubating an activated Gα alpha subunit with a GEF protein (e.g., p115 Rho GEF, Lsc, KIAA380, DrhoGEF2, or biologically-active fragments thereof, which retain GEF activity) and a monomeric G protein in the presence and absence of a test modulator to determine whether the test

30   modulator has an inhibitory, enhancing, etc. effect on the ability of an activated Gα subunit

to stimulate GEF mediated nucleotide exchange of a monomeric protein. See e.g. Hart et al., 1996, *J. Biol. Chem.* 221:25452. The test modulator can be further evaluated by conducting a second assay in which said GEF protein and a monomeric G protein, without the G protein subunit, are incubated in the presence or absence of the test modulator to determine whether

5   the test modulator had any effect on GEF mediated nucleotide exchange of the monomeric protein, and then comparing the modulation effect in the first and second assays to determine whether the modulating effect in the first assay is different from the modulating effect in the second assay, thereby indicating that the test modulator modulates the interaction of an activated Gα subunit with the GEF protein rather than the interaction of the

10   GEF protein with the monomeric G protein. For example, the stimulatory effect on GEF mediated guanine nucleotide exchange may be measured according to Example 6 below, wherein RhoA was loaded with [^{3}H]GDP and the dissociation of GDP from RhoA was measured under various conditions by the determination of bound GDP by filtration, prior to an after incubation. (See e.g. Northrup et al., *J. Biol. Chem.*, 257, 11416-11423 (1982)).

15   A method of identifying a molecule that mimics the stimulatory effect of an activated Gα subunit on GEF mediated nucleotide exchange of a monomeric G protein may also be conducted in accordance with the invention. The method comprises identifying a test compound that exhibits a binding affinity for the RGS domain of GEF proteins and then incubating a GEF protein and monomeric G protein in the presence or absence of the test

20   compound to determine whether the test compounds exhibits a stimulatory effect on GEF mediated nucleotide exchange of a monomeric G protein. The identification of test compounds that exhibit a binding affinity for the RGS domain of GEF proteins may be accomplished using techniques well known in the art. For example, an RGS polypeptide may be bound to a column and cocktails of test compounds may be passed over the column

25   to determine if any were selectively bound by the column.

A method of identifying a molecule, or mixture of molecules, that mimics the stimulatory effect of an RGS domain of GEF polypeptide on Gα subunit GTPase activity may also be conducted in accordance with the invention. The method comprises identifying a test compound that exhibits a binding affinity for a Gα subunit and incubating a GTP

30   loaded Gα subunit in the presence or absence of the test compound to determine whether the test compound exhibits a stimulatory effect GTPase activity of the Gα subunit. The

identification of test compounds that exhibit a binding affinity for the G α subunit may be
accomplished using techniques well known in the art. For example, a Gα$_{12}$ may be bound
to a substrate and incubated with both a GEF polypeptide containing an RGS domain and
the test compound to determine whether the test compound competes with the RGS domain

5   for binding to the Gα subunit.

The modulation of oncogenic transforming activity by an RGS-GEF component, or
derivatives thereof, can be measured according to various known procedures, e.g., Eva and
Aaronson, Nature, 316:273-275, 1985; Hart et al., J. Biol. Chem., 269:62-65, 1994. A
compound can be added at any time during the method (e.g., pretreatment of cells; after

10   addition of the RGS-GEF, etc.) to determine its effect on the oncogenic transforming
activity of the RGS-GEF component. Various cell lines can also be used.

Other assays for monomeric GTPase-mediated signal transduction can be
accomplished according to the invention by analogy to procedures known in the art, e.g., as
described in U.S. Pat. Nos. 5,141,851; 5,420,334; 5,436,128; and 5,482,954; WO94/16069;

15   WO93/16179; WO91/15582; WO90/00607.

The present invention thus also relates to the treatment and prevention of diseases
and pathological conditions associated with signal transduction mediated by GEF proteins
that contain an RGS domain, e.g., cancer, diseases associated with abnormal cell
proliferation. For example, the invention relates to a method of treating cancer comprising

20   administering, to a subject in need of treatment, an amount of a compound effective to treat
the disease, where the compound is a regulator of the stimulatory effect of GEF protein
containing an RGS on Gα subunit GTPase activity or where the compound is a regulator of
the stimulatory effect of a Gα subnit on GEF mediated nucleotide exchange by a monomeric
GTPase. Treating the disease can mean, delaying its onset, delaying the progression of the

25   disease, improving or delaying clinical and pathological signs of disease. A regulator
compound, or mixture of compounds, can be synthetic, naturally-occurring, or a
combination. A regulator compound can comprise amino acids, nucleotides, hydrocarbons,
lipids, polysaccharides, etc. A regulator compound is preferably a compound that regulates
expression of a GEF protein containing an RGS domain, e.g., inhibiting or increasing its

30   mRNA, protein expression, or processing, or a compound that regulates the interaction of
the RGS domain of the GEF protein with a Gα subunit. To treat the disease, the compound,

or mixture, can be formulated into pharmaceutical composition comprising a

pharmaceutically acceptable carrier and other excipients as apparent to the skilled worker.

See, e.g., *Remington's Pharmaceutical Sciences*, Eighteenth Edition, Mack Publishing

Company, 1990. Such composition can additionally contain effective amounts of other

5   compounds, especially for treatment of cancer.

EXAMPLES

Example 1. <u>Identification of homology between a Rho GEF and proteins which regulate G-
protein signaling.</u>

The RGS family of proteins act as negative regulators of G protein signalling.

10   Nineteen mammalian members of the family have been identified, all of which encode
proteins that contain a homologous core domain called the RGS box.

Examination of the sequence of p115-GEF, a GEF specific for Rho, revealed an N-
terminal region with specific homology to the conserved domain of RGS proteins, including

RGS4, RGS2, GAIP, RGS12, and RGS14 (Fig. 1). Analysis of three other Rho GEF

15   proteins, Lsc, KIAA380, and DrhoGEF also showed that they contained regions of specific

homology to the conserved domain of RGS proteins (Fig. 1).

The crystal structure of a complex between RGS4 and AlF_4-activated $G\alpha_{i1}$ revealed

that the functional core of RGS4 (the RGS box) contains nine α-helixes that fold into two

small subdomains (Tesmer et al., *Cell*, **89**, 251 (1997)). The RGS box has been shown to

20   contain the GAP activity towards $G\alpha$ subunits (Popov et al., *Proc. Natl. Acad. Sci.* USA, **94**,

7216 (1997)). The hydrophobic core residues of the box, which are conserved in members

of the RGS family, are important for stability of structure and GAP activity (Tesmer et al.,

Cell, **89**, 251 (1997) and Srinivasan et al., *J. Biol. Chem.*, **273**, 1529 (1998). RGS4

stimulates the GTPase activity of $G\alpha_{i1}$ by interacting with its three switch regions, primarily

25   by stabilization of the transition state of GTP hydrolysis (Tesmer et al., *Cell*, **89**, 251

(1997)).

Most of the hydrophobic residues that form the core of the RGS domain are

conserved in p115 Rho GEF (17 out of 23) (Fig. 1). The position of gaps in the alignment

correspond to the loops between alpha helixes of RGS domain structure. This homology

30   suggested that the N-terminal region of p115-GEF may have a similar structure to the RGS4

box domain and possess GAP activity. In contrast, the residues of RGS4 that make contact

27

with the switch regions of Gα$_{i1}$(GDP-AlF$_4$-) are not well conserved, and any GAP activity of
p115 Rho GEF will have a unique mechanism or a significantly different specificity than
those previously identified.

        A search of the gene bank revealed three other Rho-GEF members that have regions
5   homologous to the RGS region of p115. These include Lsc, KIAA380, and DrhoGEF2 (Fig.
1). Lsc appears to be the mouse homolgue of p115 RhO GEF and KIAA380 appears to be
the human homolgue of Drosophila DrhoGEF2 (Whitehead et al., *J. Biol. Chem.*, **271**,
18643 (1996); Barrett et al., *Cell*, **91**, 905 (1997)). These four Rho-GEF's define a new
· RGS related family of proteins which also possess guanine nucleotide exchange activtity for
10   Rho.

        An alignment of RGS domains of the four GEF proteins known to contain RGS
domains (p115 Rho GEF, Lsc, KIAA380, DRhoGEF2) with the RGS domains of RGS
proteins RET-RGS1, RGS1, RGS2, RGS3, RGS4, RGS7, RGS10, RGS12, RGS14,
Rap1/2B.P., and GAIP shows that a novel sub-RGS consensus sequence is defined by the
15   RGS sequence of the four GEF proteins (Fig. 18). As shown in the bottom set of sequences
shown in Fig. 18, a novel sub-RGS consensus sequence is shown by the large gap of 13 to
14 amino acids in the homology alignment, along with the conservation of amino acids on
either side of the gap.

20   Example 2. The RHO GEF protein, p115 RHO-GEF, stimulates the GTPase activity of
Gα$_{13}$ and Gα$_{12}$ subunits.

        P115 Rho GEF was tested to determine it's capability in stimulating the intrinsic
GTPase activity (GAP activity) of Gα$_{13}$ and Gα$_{12}$.

        Gα$_{12}$ was expressed in Sf9 cells and purified as described in Kozasa and Gilman, *J.
25   Biol. Chem.*, **270**, 1734 (1995). Gα$_{13}$ was prepared by a similar procedure using the
previously described baculovirus method (Singer and Miller, *J. Biol. Chem.*, **269**, 19796
(1994)) and octylglucoside during washing and elution of the α subunit after immobilization
of the heterotrimer on Ni-NTA resin (Qiagen). The eluted Gα$_{13}$ was further purified by
absorption to and elution from hydroxyapatite. Gα$_{12}$ or Gα$_{13}$ (20-30 pmol) was loaded at
30   30°C for 30 or 40 minutes, respectively with 5 μM[γ-^{32}P]GTP (50-100 cpm/fmol) and in the

presence of 5 mM EDTA. Samples were then rapidly filtered by centrifugation at 4°C

through Sephadex G50 which had been equilibrated with buffer A (50 mM NaHepes (pH

8.0), 1 mM dithiolthreitol, 5 mM EDTA, and 0.05% polyoxyethylene 10-laurylether) to

remove free $[\gamma\text{-}^{32}P]GTP$ and $[^{32}Pi]$. Hydrolysis of GTP was initiated by adding $G\alpha$ loaded

5   with $[\gamma\text{-}^{32}P]GTP$ in buffer A containing 8 mM MgSO$_4$, 1 mM GTP and the indicated amount

of p115. The reaction mixture was incubated at 4°C or 15°C. Aliquots (50 µl) were

removed at the indicated times and mixed with 750 µl of 5%(w/v) NoritA in 50 mM

NaH$_2$PO$_4$. The mixture was centrifuged at 2000 rpm for 5 minutes and 400 µl of

supernatant containing ^{32}Pi were counted by liquid scintillation spectrometry.

10   The hydrolysis of GTP bound to $G\alpha_{13}$ and $G\alpha_{12}$ was performed at 15°C either with

or without 10nM full-length p115 (Fig. 2, Panel A). The hydrolysis of GTP bound $G\alpha_{13}$ and

$G\alpha_{12}$ was measured at 4°C in the presence of various concentrations of p115 (Fig. 2, Panel

B). Full-length p115 was able to stimulate a single round of hydrolysis of $[\gamma\text{-}^{32}P]GTP$ which

had been prebound to the $G\alpha_{13}$ subunit. The intrinsic GTPase activity of $G\alpha_{12}$, the closest

15   homologue of $G\alpha_{13}$, was also stimulated by full-length p115. At 15°C, the k$_{cat}$ for

hydrolysis of GTP by $G\alpha_{12}$ (0.07 min^{-1}) and $G\alpha_{13}$ (0.24 min^{-1}) were respectively increased

5-fold and 10-fold by 10 nM p115 (Fig. 2, Panel A). Similar results were obtained with

several preparations of $G\alpha_{12}$ and $G\alpha_{13}$. Treatment of p115 at 90°C inactivated this GAP

activity. Due to the rapid hydrolytic rates of $G\alpha_{13}$, assays were performed at 4°C to better

20   estimate the effect of p115 on the initial rate of GTPase activity by the G protein (Fig. 2,

Panel B). Under these conditions, 100 nM p115 caused and 80-fold increase in the GTPase

activity of $G\alpha_{13}$. In contrast, the hydrolytic rate of $G\alpha_{12}$ was increased only 6-fold.

Although stimulation of both proteins was observed at concentrations of p115 as low as 1

nM, measurements at both temperatures indicate that p115 is a more efficacious GAP for

25    $G\alpha_{13}$ than $G\alpha_{12}$.

In the absence of a receptor, the rate limiting step in the binding of GTPγS to $G\alpha$ and

the steady state hydrolysis of GTP is the release of GDP. P115 did not affect either the rate

of GTPγS binding to $G\alpha_{12}$ and $G\alpha_{13}$ or the steady state of GTPase activity of either subunit.

29

Therefore, p115 stimulates only the intrinsic GTPase activity of $G\alpha_{12}$ and $G\alpha_{13}$ without effecting their rates of nucleotide exchange.

The conserved RGS box region of RGS proteins is sufficient to show GAP activity *in vitro* (Popov et al., *Proc. Natl. Acad. Sci.* USA, **94**, 7216 (1997)). Therefore, a fusion

5   protein (Fig. 1, Panel B) of glutathione-S-transferase and the N-terminal region of p115, GST-RGS, was tested for GAP activity. This region retains RGS homology domain but not the Dbl or PH domains of p115. This "RGS domain" of p115 (10 nM) was almost as active as full-length p115 when tested for GAP activity for $G\alpha_{12}$ and $G\alpha_{13}$ (Fig. 3). In contrast, a construct of p115 missing this N-terminal region was ineffective. Thus, the data indicates

10   that the RGS homology region is responsible for the GAP activity of p115.

Example 3. The p115 RHO-GEF, does not stimulate the GTPase activity of $G\alpha_i$, $G\alpha_z$ $G\alpha_q$ and $G\alpha_s$ subunits.

The specificity of the GAP activity of p115 for various G protein α subunits was

15   examined as follows.

$G\alpha_S$ was expressed in and purified from *Escherichia coli* as described in Lee et al., *Meth. Enzymol.*, **237**, 146 (1994). $G\alpha_i$, $G\alpha_z$, and $G\alpha_qR183C$ were expressed in Sf9 cells and purified as described in Kozasa and Gilman, *J. Biol. Chem.*, **270**, 1734 (1995) and Biddlecome et al., *J. Biol. Chem.*, **271**, 7999 (1996). $G\alpha_i$, $G\alpha_z$, and $G\alpha_z$ were loaded with 5-

20   10 μM [γ-^{32}P]GTP at 20°C (for $G\alpha_s$) or 30°C (for $G\alpha_i$ and $G\alpha_z$) for 20 minutes in the presence of 5 mM EDTA and GAP assays were performed as described above for $G\alpha_{12}$ and $G\alpha_{13}$. Gap activity on $G\alpha_q$ was assessed with a mutant $G\alpha_qR183C$. An analogous mutation in $G\alpha_i$ R1178C, causes markedly reduced GTPase activity but response to RGS proteins was retained (Berman et al., *Cell*, **86**, 445 (1996)). The slow GTPase activity of $G\alpha_qR183C$

25   enables loading of [γ-^{32}P]GTP on $G\alpha_q$ without using receptor to accelerate nucleotide exchange. $G\alpha_qR183C$ was loaded with 10 μM [γ-^{32}P]GTP in the presence of 50 mM Hepes (pH7.4), 0.1 mg/ml BSA, 1 mM DTT, 1 mM EDTA, 0.9 mM MgSO$_4$, 30 mM (NH$_4$)$_2$SO$_4$, 4% glycerol, and 5.5 mM CHAPS at 20°C for 2 hours. The reaction mixture was rapidly

filtered through Sephadex G50 which had been equilibrated with 50 mM Hepes (pH 7.4), 1 mM DTT, 1 mM EDTA, 0.9 mM SO_4, 0.1 mg/ml BSA, and 1 mM CHAPS.

The results of this study showed that p115 (100 nM) did not stimulate the GTPase activity of $G\alpha_i$, $G\alpha_z$, or $G\alpha_q$ under conditions where RGS4 acts as a GAP for these $G\alpha$

5   subunits (Figure 4). Similarly, p115 did not accelerate the GTPase activity of $G\alpha_s$, nor did p115 Rho GEF have any GAP activity towards RhoA or rac1. Thus, p115 is a GAP with specificity for $G\alpha_{12}$ and $G\alpha_{13}$.

Example 4. <u>Selective inhibition of p115 GAP activity by AlF_4^- activated forms of $G\alpha$</u>

10 <u>subunits.</u>

RGS proteins have been shown to have high affinity for the GDP-AlF_4^- bound form of α subunits, a configuration similar to the transition state of GTP hydrolysis (Tesmer et al., *Cell*, **89**, 251 (1997), Berman et al., *J. Biol. Chem.*, **271**, 27209 (1996)). Therefore, the GDP-AlF_4^- forms of $G\alpha$ should compete with $G\alpha$GTP for interaction with p115 and block

15 the observed GAP activity. As shown in Fig. 5, Panel A, GDP-AlF_4^- bound $G\alpha_{12}$ and $G\alpha_{13}$ effectively inhibited the GAP activity of p115 for $G\alpha_{12}$, while similar forms of $G\alpha_s$, $G\alpha_i$, and $G\alpha_q$ were without effect. Additionally, a tritration of GDP-AlF_4^- bound forms of $G\alpha_{12}$ and $G\alpha_{13}$ demonstrated that the subunits are equipotent in inhibiting the GAP activity of $G\alpha_{13}$ (Fig. 5, Panel B). These competition assays suggest that the two G protein subunits

20 have a similar affinity for p115 and supports the apparent differential efficacy of p115 towards the subunits as shown in Fig. 2.

Example 5. <u>Binding of $G\alpha_{13}$ to p115 Rho GEF in vivo.</u>

The following experiments demonstrated that $G\alpha_{13}$ and p115 Rho GEF interact in a

25 GTP-dependent manner.

EXV-myc tagged (for COS cell transfections) and pAc-Glu tagged (for baculovirus expression) proteins with deletions of the RGS or DH domains were constructed as previously described in Hart et al., *J. Biol. Chem.*, **271**, 25452-25458 (1996). Full-length versions were constructed in the same vectors. A fusion of GST to the first 246 amino acids

30 of p115 Rho GEF was constructed in pGEX4T-2 (Pharmacia). Transfections,

31

immunoprecipitations, and purifications were performed as previously described in Hart et al., *J. Biol. Chem.*, **271**, 25452-25458 (1996).

In COS cells transfected with myc-tagged p115 Rho GEF, $G\alpha_{13}$ can be specifically immunoprecipitated using the anti-myc antibody (Fig. 6, Panels A and B). This interaction

5   is dependent on the presence of aluminum fluoride which is added to mimic the activated GTP-bound state of the $G\alpha_{13}$. Additionally, a truncated mutant of p115 Rho GEF which lacks the amino-terminal RGS domain is incapable of mediating co-immunoprecipitation, while full-length protein with a deletion in the DH domain does mediate co-immunoprecipitation. The differential binding of full-length and truncated Rho GEF

10 .   proteins could also be detected using antibodies to $G\alpha_{13}$ to immunoprecipitate the complex (Fig. 6, Panel C). A very weak interaction with $G\alpha_{12}$ was detectable, while antibodies to $G\alpha_s$, $G\alpha_i$, $G\alpha_q$ and $G\alpha_z$ do not detect immunoreactive bands in the anti-myc immunoprecipitates, in spite of the fact that their respective antigens are detectable in the whole cell lysates. The co-immunoprecipitation of p115 Rho GEF and $G\alpha_{13}$ can be

15   reproduced in a semi-purified system in which purified $G\alpha_{13}$ is added to immunoprecipitated p115 Rho GEF (Fig. 6, Panel D), suggesting a direct interaction. This direct interaction is consistent with the observation that p115 Rho GEF stimulates $G\alpha_{13}$ GTPase activity, but . also indicates that p115 Rho may be an effector of $G\alpha_{13}$.

Binding could also be detected between the Rho GEF protein, KIAA380 and the α_{12}

20   G protein subunit (Fig. 9, KIAA380 is referred to as FL147). In COS cells transfected with myc-tagged KIAA380, $G\alpha_{12}$ can be specifically immunoprecipitated using the anti-myc antibody (Fig. 9. Panels A and B, KIAA380 is referred to as FL147). This interaction is dependent on the presence of aluminum fluoride which is added to mimic the activated GTP-bound state of the $G\alpha_{13}$.

25

Example 6.   Stimulation of p115 Rho GEF activity by $G\alpha_{13}$.

The ability of $G\alpha_{13}$ to affect the exchange activity of p115 Rho GEF was examined by incubating RhoA and p115 Rho GEF with or without $G\alpha_{13}$ to determine the effect on guanine nucleotide exchange.

RhoA (2.5 µM) was loaded with [³H]GDP by incubation at 30°C for 1 hour with 25

µM GDP (10,000 cpm/pmol) in 50 mM NaHepes, pH 7.5, 50 mM NaCl, 4 mM EDTA,

1mM dithiolthreitol and 0.1%Triton X-100. After addition of $MgCl_2$ to 9 mM and

octylglucoside to 1%, the Rho was incubated for an additional 5 minutes and separated from

5   free GDP by rapid filtration through Sephadex-G50 that had been equilibrated with 50 mM

NaHEPES, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM dithiolthreitol, 5 mM $MgCl_2$, and 1%

octylglucoside. Dissociation of GDP from RhoA was measured at 30°C in 20 µl of 50 mM

NaHEPES, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM dithiolthreitol, 5 mM $MgCl_2$, 30 mM

$AlCl_3$, 5 mM NaF, and 5 µM GTPγS. Unless specified, G protein alpha subunits were

10   preincubated with AMF (30 µM $AlCl_3$, 5 mM $MgCl_2$ and 5 mM NaF) prior to mixing with

other proteins. Where indicated, alpha subunits were treated with 25 µM GTPγS or GDPβS

rather than AMF and reactions were incubated without AMF but with 5 µM of the

respective nucleotide. Reactions were started with the addition of [³H]-GDP-RhoA and

bound GDP was determined by filtration (Northup et al., *J. Biol. Chem.*, **257**, 11416-11423

15   (1982)) prior to and after incubation.

The $G\alpha_s$ and $G\alpha_i$ alpha subunits were purified after expression in *Escherichia coli*

(Lee et al., *Meth. Enzymol.*, **237**, 146-164 (1994)). The $G\alpha_q$ and $G\alpha_z$ alpha subunits were

coexpressed in Sf9 cells with hexahistidine-tagged beta and gamma subunits and isolated as

described (Kozasa and Gilman, *J. Biol. Chem.*, **270**, 1734-1741 (1995)). $G\alpha_{13}$ was prepared

20   by a similar procedure to $G\alpha_{12}$ using baculovirus (Singer et al., *J. Biol. Chem.*, **269**, 19796-

19802 (1994)) and octylglucoside during washing and elution of the α subunit after

immobilization of the heterotrimer on Ni-NTA resin (Qiagen). The eluted $G\alpha_{13}$ was further

purified by absorption to and elution from hydroxyapatite. About 500 ug of purified $G\alpha_{13}$

can be obtained from 3 liters of cells. The expression of GST-RhoA in Sf9 cells, cleavage

25   of the GST tag and isolation of the free RhoA were as described in Singer et al., *J. Biol.*

Chem., **271**, 4505-4510, (1996).

These studies demonstrated that the $G\alpha_{13}$ is capable of stimulating the activity of

full-length p115 Rho GEF in a manner which depends on the concentrations of both p115

Rho GEF (Fig. 7, Panel A) and $G\alpha_{13}$ (Fig. 7, Panel B). The closely related alpha subunit

$G\alpha_{12}$ was ineffective in stimulating the activity of p115 Rho GEF in these experiments (Fig.

7, Panel A). Stimulation of Rho exchange was also monitored as a function of the activation

state of $G\alpha_{13}$. The data graphed in Fig. 7, Panel C confirm that the stimulation of exchange

activity is dependent on either aluminum fluoride (AMF) or GTPγS, but is not stimulated by

5   the deactivated nucleotide state mimicked by GDPβS. Additionally, a series of other alpha

subunits including $G\alpha_q$, $G\alpha_z$, $G\alpha_s$, and $G\alpha_i$ also did not affect the activity of p115 Rho GEF

(Fig. 7, Panel D). These results are consistent with with the activated $G\alpha_{13}$ -dependent

binding shown in Figure 6, and suggest that the productive binding of $G\alpha_{13}$ to p115 Rho

GEF may be sufficient for activation.

10

Example 7. <u>Effects of domains of p115 and $G\alpha_{12}$ on the p115 nucleotide exchange activity</u>.

The theory that the RGS domain of p115 Rho GEF is normally autoinhibitory and

that binding to $G\alpha_{13}$ relieves this inhibition was examined by comparing the effects of full-

length Rho-GEF versus truncated Rho-GEF on Rho exchange activity.

15   Preparation of p115 proteins was as described in Example 1 above and as described

in Hart et al., *J. Biol. Chem.*, **271**, 25452-25458 (1996). The assays shown in Figure 8,

Panels B and C were performed as described in Example 5 above. AMF was the activating

agent.

The results of these experiments showed that truncated p115 Rho GEF lacking the

20   RGS domain demonstrates consistently elevated Rho exchange activity when compared with

equal concentrations of the full-length protein (Fig. 8, Panel A). Additionally, addition of

the isolated RGS domain (as a GST fusion protein) resulted in abrogation of $G\alpha_{13}$-

stimulated p115 Rho GEF activity (Fig. 8, Panel B). These data do not preclude additional

$G\alpha_{13}$-binding sites on p115 Rho GEF, although they do suggest a primary mode of action

25   via the RGS domain.

The inability of the $G\alpha_{12}$ subunit to activate p115 Rho GEF was puzzling in light of

the fact that p115 Rho GEF is capable of activating the GTPase of both $G\alpha_{12}$ and $G\alpha_{13}$.

Therefore, an experiment was conducted in which $G\alpha_{12}$ was added to a $G\alpha_{13}$-stimulated

p115 Rho GEF assay (Fig. 8, Panel C). The results showed that $G\alpha_{12}$ was able to inhibit the

34

coupling of $G\alpha_{13}$ with p115 Rho GEF. This data is consistent with a model in which $G\alpha_{12}$ competes with $G\alpha_{13}$ for binding to the RGS domain of p115 Rho GEF. However, binding of $G\alpha_{12}$ to p115 Rho GEF is clearly not sufficient to stimulate Rho exchange activity. These results suggest that either the interaction of $G\alpha_{12}$ with the RGS domain of p115 Rho GEF is

5   quite different from that of $G\alpha_{13}$ or that there may be an additional site of interaction between $G\alpha_{13}$ and p115 Rho GEF.

For other aspects of the nucleic acids, polypeptides, antibodies, etc., reference is made to standard textbooks of molecular biology, protein science, and immunology. See, e.g., Davis et al. (1986), *Basic Methods in Molecular Biology*, Elsevir Sciences Publishing,

10   Inc., New York; Hames et al. (1985), *Nucleic Acid Hybridization*, IL Press, *Molecular Cloning*, Sambrook et al.; *Current Protocols in Molecular Biology*, Edited by F.M. Ausubel et al., John Wiley & Sons, Inc; *Current Protocols in Human Genetics*, Edited by Nicholas C. Dracopoli et al., John Wiley & Sons, Inc.; *Current Protocols in Protein Science*; Edited by John E. Coligan et al., John Wiley & Sons, Inc.; *Current Protocols in Immunology*; Edited

15   by John E. Coligan et al., John Wiley & Sons, Inc. The entire disclosure of all patent applications, patents, and publications cited herein are hereby incorporated by reference.

From the foregoing description, on skilled in the art can easily ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and

20   conditions.

CLAIMS

What is claimed is:

1. An isolated RGS-GEF polypeptide, or a biologically active fragment thereof,
   consisting essentially of an RGS domain of a GEF protein.

2. An isolated RGS-GEF polypeptide, or a biologically active fragment thereof,
   comprising an RGS domain of a GEF protein, with the proviso that the polypeptide
   does not comprise a DH domain or a PH domain.

3. An isolated RGS-GEF polypeptide, or a biologically active fragment thereof,
   wherein the polypeptide is selected from the group consisting of p115 Rho-GEF,
   Lsc, KIAA380, and wherein the polypeptide is mutated in the RGS domain, and
   wherein the polypeptide has a specific binding affinity for a G protein α subunit or a
   GTPase activating activity for a G protein α subunit.

4. An isolated RGS-GEF polypeptide, or a biologically active fragment thereof,
   according to claim 1 or 2, wherein the GEF protein is a Rho GEF protein.

5. An isolated RGS-GEF polypeptide, or a biologically active fragment thereof,
   according to claim 4, wherein the Rho GEF protein is p115 Rho-GEF.

6. An isolated RGS-GEF polypeptide, or biologically active fragment thereof,
   according to claim 4 wherein the Rho GEF protein is selected from the group
   consisting of Lsc, KIAA380, and DrhoGEF2.

7. An isolated RGS-GEF polypeptide, or a biologically active fragment thereof,
   according to claim 1 or 2, wherein the polypeptide has a specific binding affinity for
   a G protein α subunit or a GTPase activating activity for G protein α subunits.

8.   An isolated RGS-GEF polypeptide, or a biologically active fragment thereof, according to claim 4, wherein the polypeptide has a specific binding affinity for a G protein α subunit or a GTPase activating activity for G protein α subunits.

5   9.   An isolated RGS-GEF polypeptide, or a biologically active fragment thereof, according to claim 5, wherein the polypeptide has a specific binding affinity for a G protein α subunit or a GTPase activating activity for G protein α subunits.

10.   An isolated RGS-GEF nucleic acid consisting essentially of a nucleotide sequence

10   encoding a polypeptide comprising an RGS domain of a GEF protein.

11.   An isolated RGS-GEF nucleic acid comprising a nucleotide sequence encoding a polypeptide comprising an RGS domain of a GEF protein, wherein the polypeptide does not include a DH domain or a PH domain.

15

12.   An isolated RGS-GEF nucleic acid according to claim 10 or 11, wherein the GEF protein is a Rho GEF protein.

13.   An isolated RGS-GEF nucleic acid according to claim 12, wherein the Rho GEF

20   protein is p115 Rho GEF.

14.   An isolated RGS-GEF nucleic acid according to claim 12 wherein the Rho GEF protein is selected from the group consisting of Lsc, KIAA380, and DrhoGEF2.

25   15.   An isolated RGS-GEF nucleic acid according to claim 10 or 11, wherein the polypeptide has a specific binding affinity for a G protein α subunit or a GTPase activating activity for a G protein α subunit.

16.   An isolated RGS-GEF nucleic acid according to claim 12, wherein the polypeptide

30   has a specific binding affinity for a G protein α subunit or a GTPase activating activity for a G protein α subunit.

37

17.   An isolated RGS-GEF nucleic acid according to claim 13, wherein the polypeptide
       has a specific binding affinity for a G protein α subunit or a GTPase activating
       activity for a G protein α subunit.

5

18.   A method of modulating an activity of a G protein α subunit comprising,
       administering to a mammal an effective amount of a polypeptide according to claim
       1 or 4.

10   19.   A method of identifying or assaying a molecule that inhibits or enhances binding of a
       monomeric G protein guanine nucleotide exchange factor to a G protein α subunit
       comprising incubating the G protein α alpha subunit, or fragments thereof, with the
       monomeric G protein nucleotide exchange factor, or fragments thereof, in the
       presence and absence of a test molecule and determining whether the presence of the
15   test molecule inhibits or enhances binding between the monomeric G-protein
       guanine nucleotide exchange factor and the G protein α subunit.

20.   A method of identifying or assaying a molecule that inhibits or enhances a
       stimulatory effect of a GEF on a Gα subunit GTPase activity comprising incubating
20   a Gα alpha subunit, or fragments thereof, with a GEF protein, or fragments thereof,
       in the presence and absence of a test molecule and determining whether the presence
       of the test molecule inhibits or enhances the stimulatory effect of the GEF protein on
       Gα subunit GTPase activity.

25   21.   A method of identifying or assaying a molecule that specifically inhibits the
       stimulatory effect of an activated Gα subunit on GEF mediated nucleotide exchange
       of a monomeric G protein, compising conducting a first assay by incubating an
       activated Gα alpha subunit, or fragments thereof, with a GEF protein, or fragments
       thereof, and a monomeric G protein, or fragments thereof, in the presence and
30   absence of a test inhibitor, conducting a second assay by incubating a GEF protein,

38

or fragments thereof, and a monomeric G protein, or fragments thereof, in the presence and absence of the test inhibitor, and determining whether any inhibitory effect of the test inhibitor in the first assay is greater than any inhibitory effect of the test inhibitor in the second assay.

5

22. A method of identifying or assaying a molecule that specifically enhances the stimulatory effect of an activated $G\alpha$ subunit on GEF mediated nucleotide exchange of a monomeric G protein, compising conducting a first assay by incubating an activated $G\alpha$ alpha subunit, or fragments thereof, with a GEF protein, and fragments

10   thereof, and a monomeric G protein, or fragments thereof, in the presence and absence of a test enhancer, conducting a second assay by incubating a GEF protein, or fragments thereof, and a monomeric G protein, or fragments thereof, in the presence and absence of the test enhancer, and determining whether any enhancing effect of the test enhancer in the first assay is greater than any enhancing effect of the

15   test enhancer in the second assay.

23. A method of identifying or assaying a molecule that mimics the stimulatory effect of an activated $G\alpha$ subunit on GEF mediated nucleotide exchange of a monomeric G protein comprising identifying a test compound that exhibits a binding affinity for

20   the RGS domain of GEF proteins, or fragments thereof, incubating a GEF protein,or fragments thereof, and monomeric G protein, or fragments thereof, in the presence or absence of the test compound, determining whether the test compound exhibits a stimulatory effect on GEF mediated nucleotide exchange of a monomeric G protein.

25   24. A method of identifying or assaying a molecule that mimics the stimulatory effect of an RGS domain of a GEF protein on GTPase activity of a $G\alpha$ subunit comprising identifying a test compound that exhibits a binding affinity for a $G\alpha$ subunit and incubating a GTP loaded $G\alpha$ subunit in the presence or absence of the test compound to determine whether the test compound has a stimulatory effet on $G\alpha$

30   subunit GTPase activity.

39

25.   A method according to claim 19, 20, 21, 22, 23, or 24 wherein the GEF protein is selected from the group consisting of p115 Rho GEF, Lsc, KIAA380, and DrhoGEF2.

5

26.   A method of expressing in transformed host cells, a polypeptide coded for by a nucleic acid, comprising culturing transformed host cells containing a nucleic acid according to claim 11.

10   27.   A transformed cell containing a nucleic acid according to claim 11.

28.   A vector comprising a nucleic acid according to claim 11.

```
                                        * * *    * *
RGS4 2nd St.  aaaaaaaaaaaaaaaaaaaaaaaaaa---..aaaaaaaaaaaaaa..
RGS4    (59)  WAESLENLINHECGLAAFKAFLKSE---YSEENIDFWISCEEYKKI
RGS2    (80)  WAEAFDELLASKYGLAAFRAFLKSE---FCEENIEFWLACEDFKKT
GAIP    (87)  WAQSFDKLMHSPAGRSVFRAFLRTE---YSEENMLFWLACEELKAE
RGS12  (712)  WAVSFERLLQDPVGVRYFSDFLRKE---FSEENILFWQACECFSHV
RGS14   (64)  WAQSFERLLQDPRGLAYFTEFLKKE---FSAENVTFWQACERFQQI
p115    (45)  QFQSLEQVKRRPAHLMALLQHVALQ---FEPGPLLCCLHADMLGSL
Lsc     (43)  QFQSLEQVKRRPAHLMALLQHVALQ---FEPGPLLCCLHADMLSSL
KIAA380(310)  IFQDLEKLKSRPAHLGVFLRYIFSQ---ADPSPLLFYLCAEVYQQA
DrhoGEF2(924) PFNNLTRLL-EAENVTFLAIFLNYVISNSDPAPLLFYLITELYKEG

                                                    * *   *
RGS4 2nd St.  --....aaaaaaaaaaaaa..-----....----..aaaaaaa.
RGS4   (102)  --KSPSKLSPKAKKIYNEFISV-----QATKE----VNLDSCTREE
RGS2   (123)  --KSPQKLSSKARKIYTDFIEK-----EAPKE----INIDFQTKSL
GAIP   (130)  --ANQHVVDEKARLIYEDYVSI-----LSPKE----VSLDSRVREG
RGS12  (755)  PAHDKKELSYRAREIPSKFLCS-----KATTP----VNIDSQAQLA
RGS14  (207)  PASDTKQLAQEAHNIYHEFLSS-----QALSP----VNIDRQAWLS
p115    (88)  ---GPKEAKKAFLDFYHSFLEKTAVLR-VPVPPNVAFELDRTRADL
Lsc     (86)  ---GPKEAKKAFLDFYHSFLEKTAVLR-VPVPPSVAFELDRTRPDL
KIAA380(353)  ---SPKDSRSLGKDIWNIFLEKNAPLR-VKIPEMLQAEIDSRLRN-
DrhoGEF2(969) ---TSKDMRKWAYEIHSTFLVPRAPLSWYRQDESLAREVDNVLQLE

                                *     * *    *
RGS4 2nd St.  aaaaa..--.....aaaaaaaaaaaaaaaaaaaaaaaaaaaa
RGS4   (137)  TSRNMLE--PTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLD (176)
RGS2   (158)  IAQNIQE--ATSGCFTTAQKRVYSLMENNSYPRFLESEFYQD (197)
GAIP   (165)  INKKMQE--PSAHTFDDAQLQIYTLMHRDSYPRFLSSPTYRA (204)
RGS12  (792)  DDILNA---PHPDMFKEQQLQIFNLMKFDSYTRFLKSQLYQE (830)
RGS14  (144)  EEVLAQ---PRPDMFRAQQLQIFNLMKFDSYARFVKSPLYQE (182)
p115   (130)  ISEDVQ-RRFVQEVVQSQQVAVGRQLEDFRSKRLMGMTPWEQ (170)
Lsc    (128)  ISEDVQ-RRFIQEVVQSQQAAVSRQLEDFRSKRLMGMTPWEQ (168)
KIAA380(394)  -SEDA--RGVLCEAQEAAMPEIQEQIHDYRTKRTLGLGSLYG (432)
DrhoGEF2(1012)YDKVEILRTVFLRSRKRAKDLISEQLREFQQKRTAGLGTIYG (1053)
```
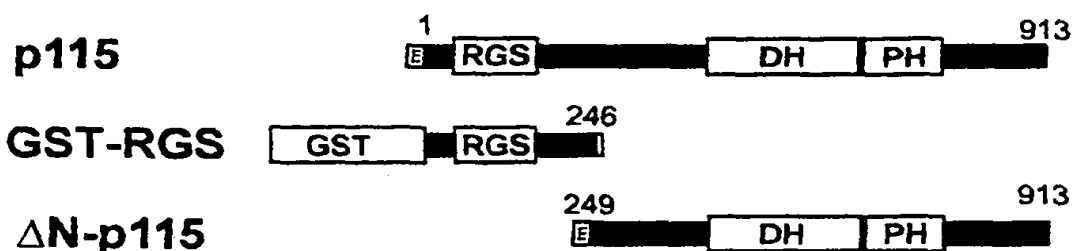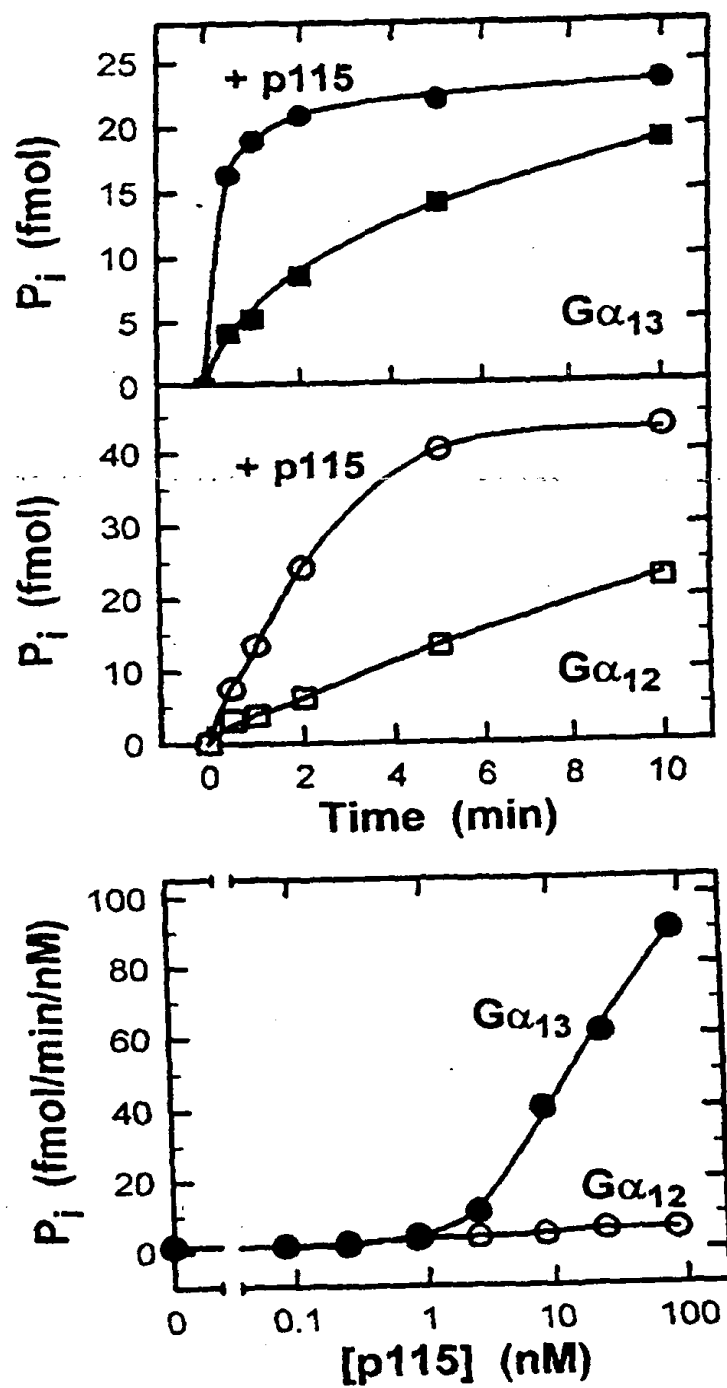
p115

GST-RGS

ΔN-p115

FIGURE 1

2/34

FIGURE 2
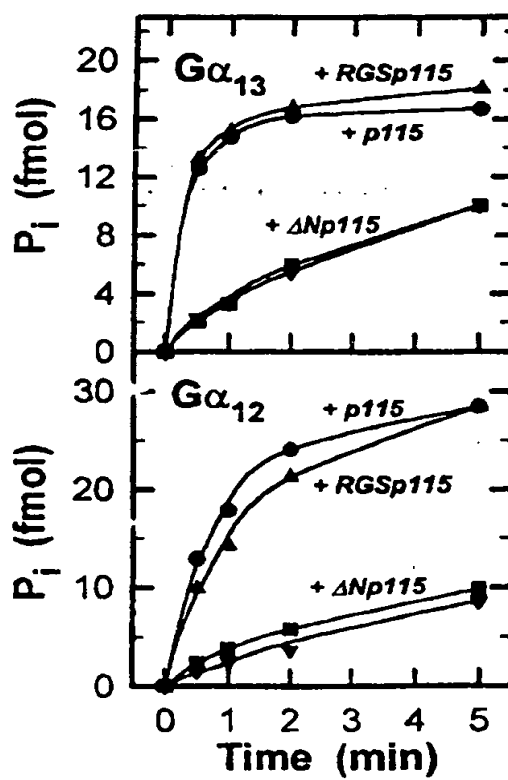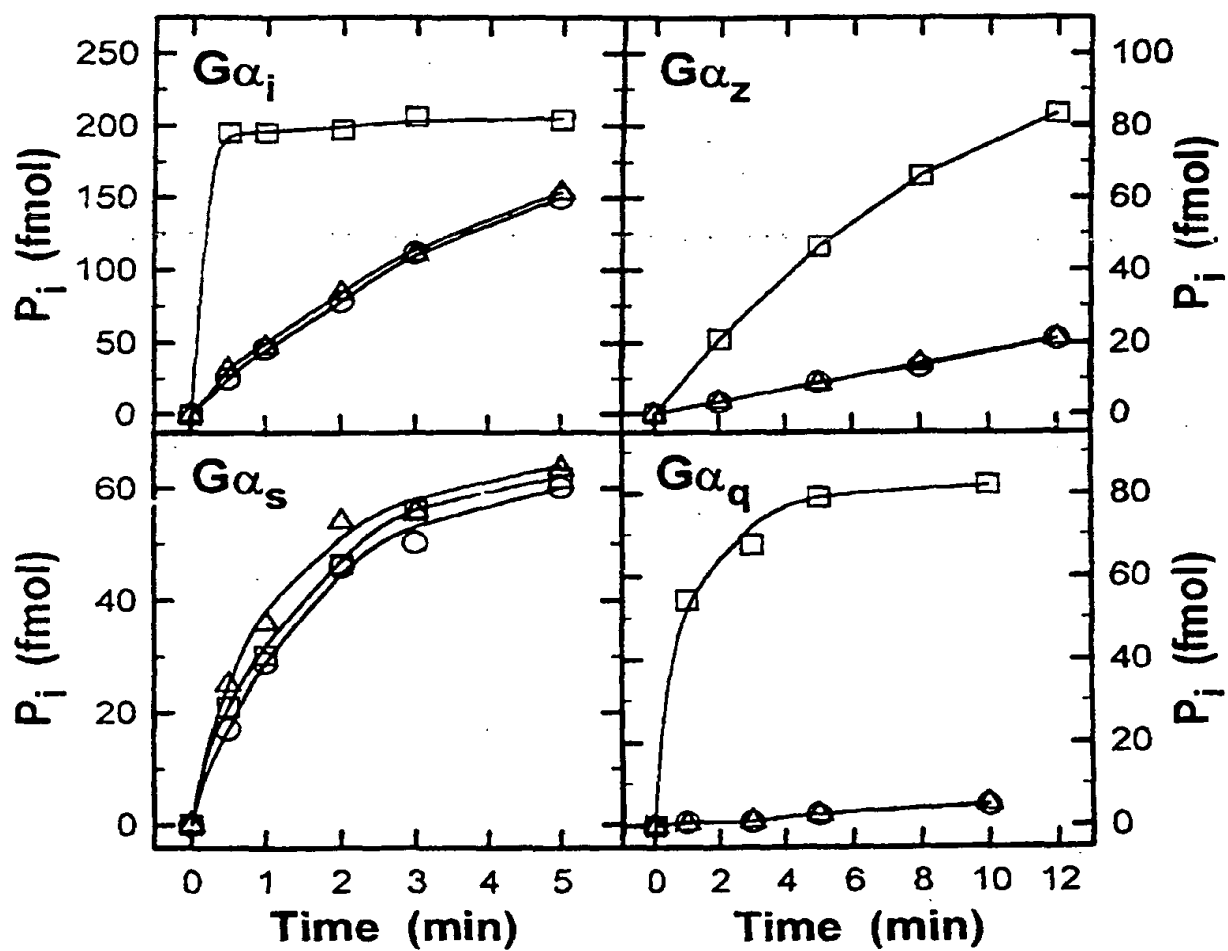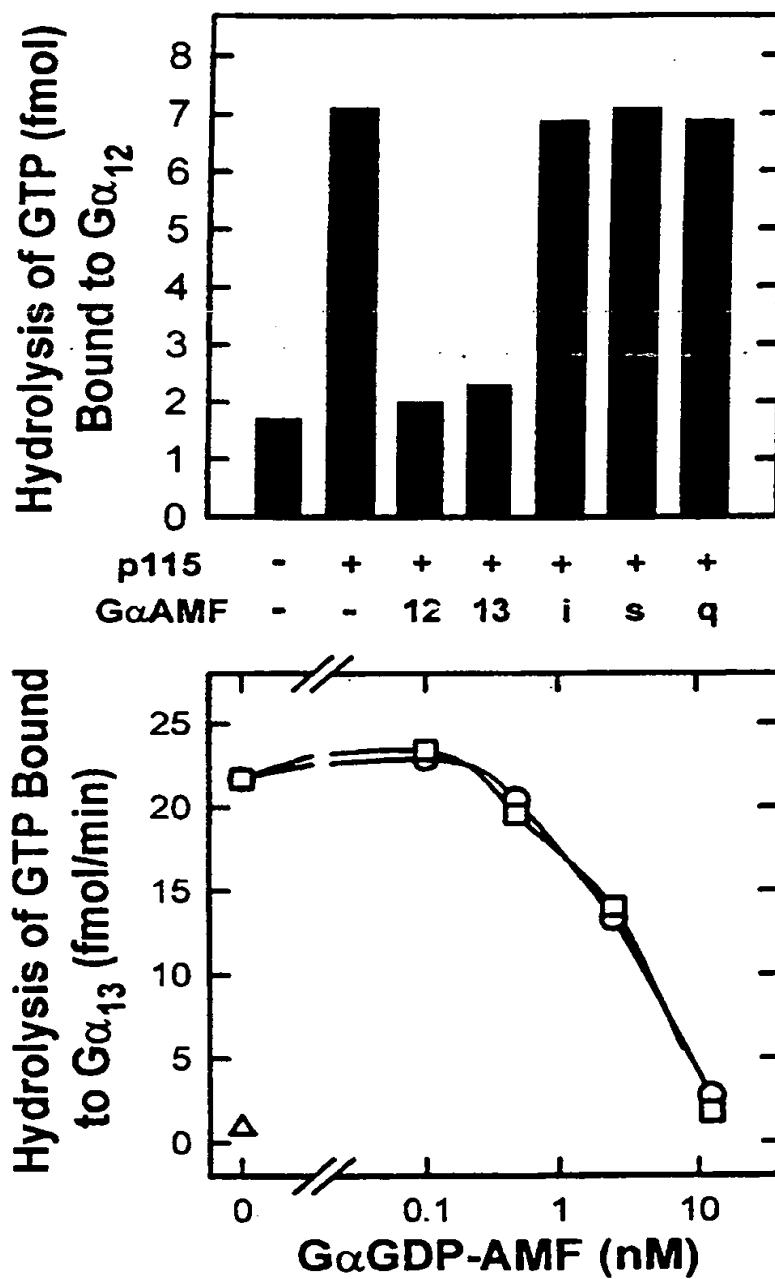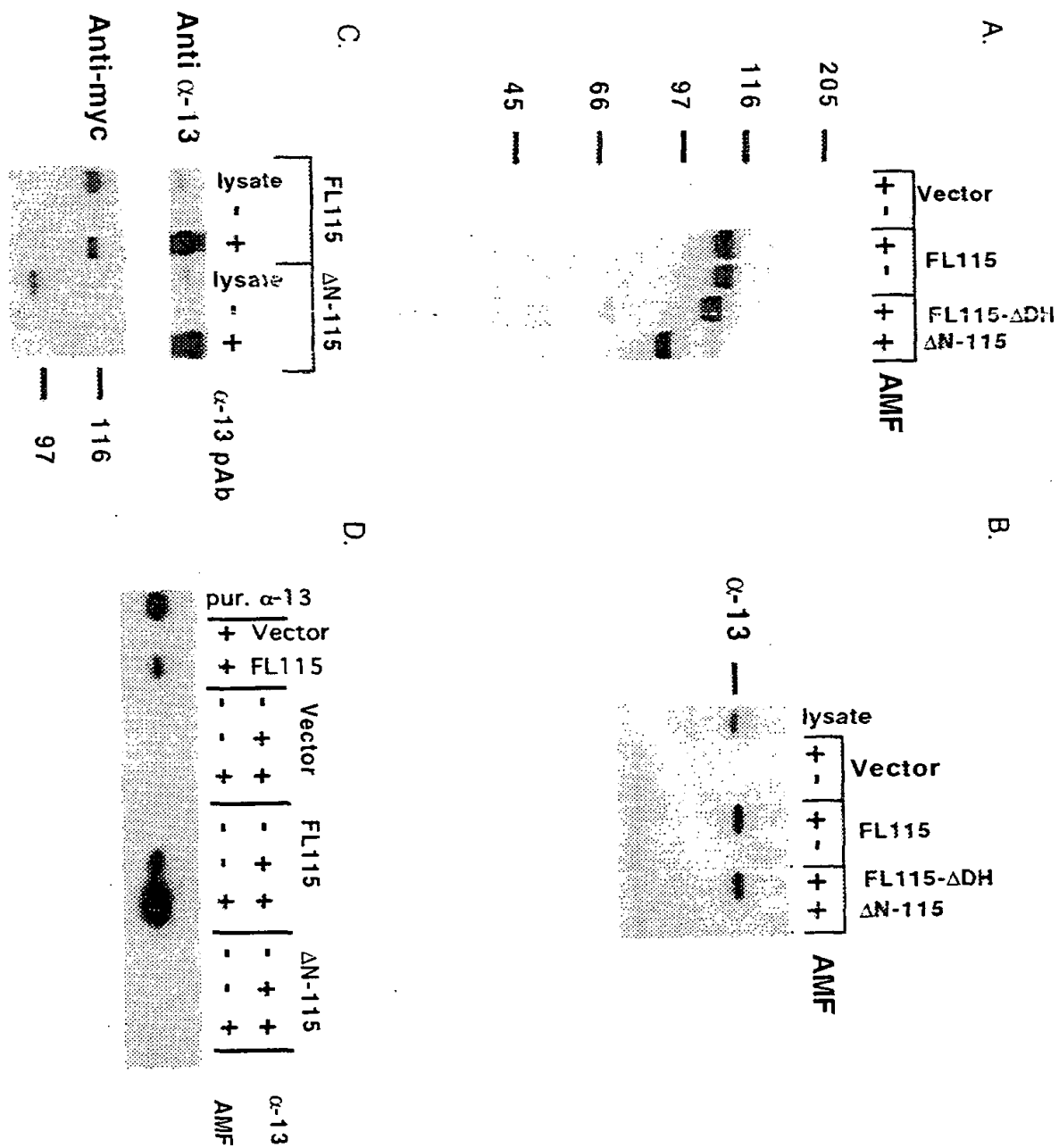
FIGURE 3

FIGURE 4

FIGURE 5

FIGURE 6

FIGURE 7

FIGURE 8

A. anti-myc   **B. anti-α-12**
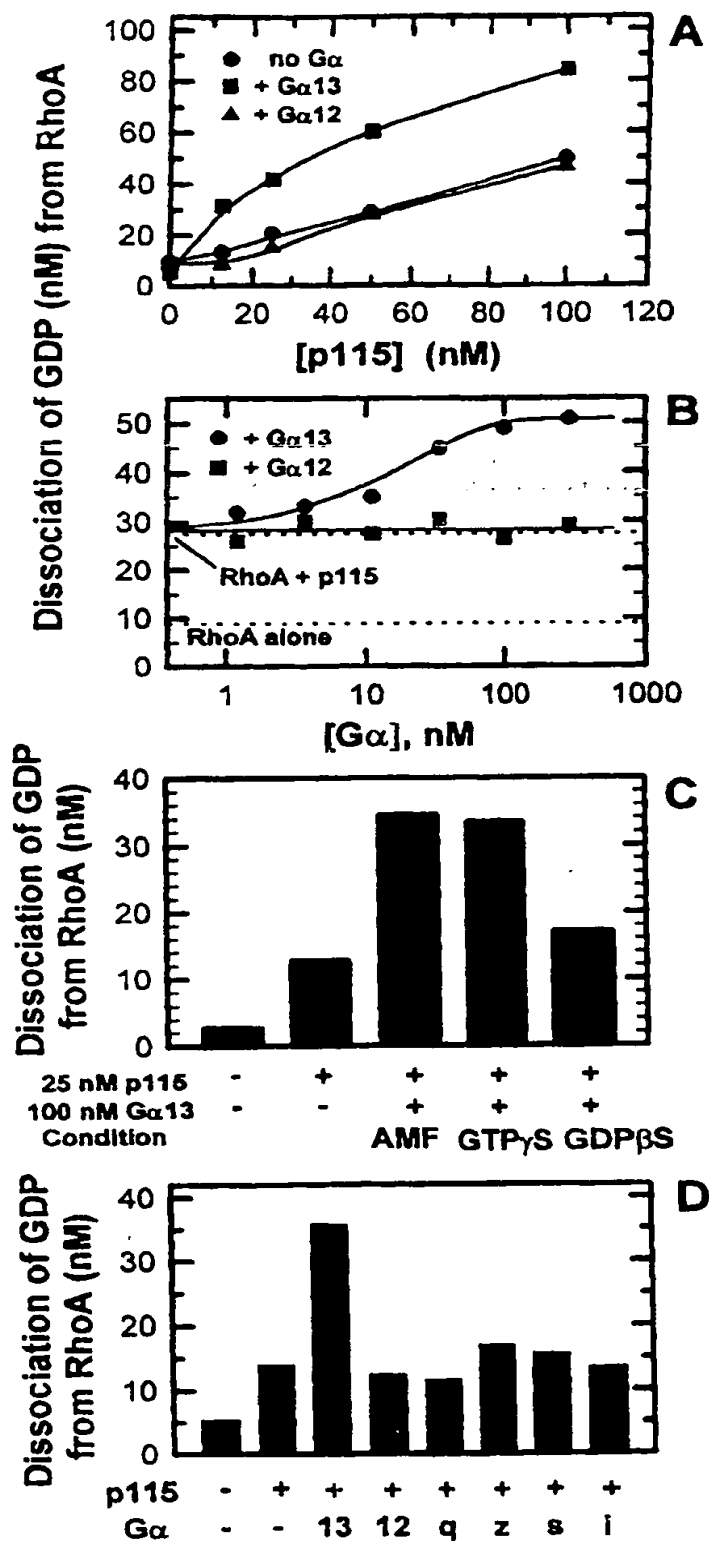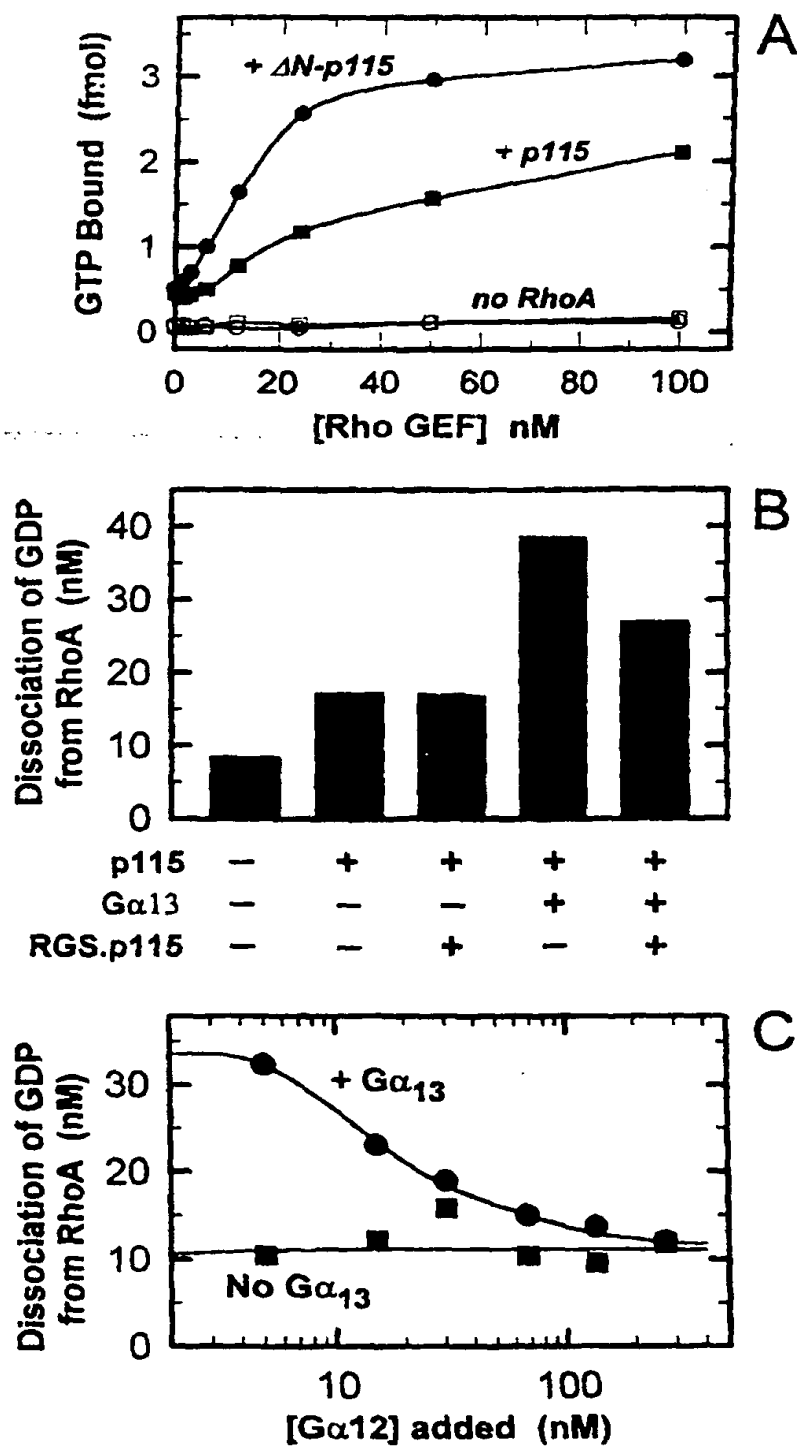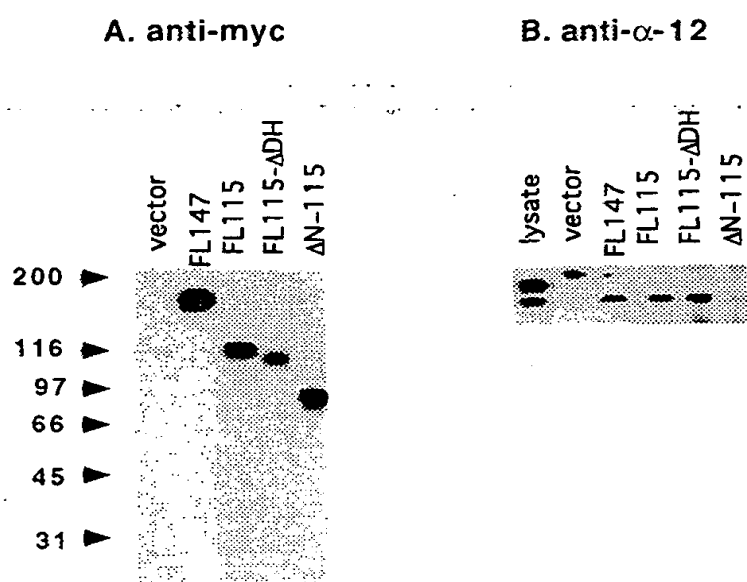
FIGURE 9

```
              10        20        30        40
    ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴_____

medfargaaspgpsrpglvpvsiigaededfeneletnse 40
eqnsqfqsleqvkrrpahlmallqhvalqfepgpllcclh 80
admlgslgpkeakkafldfyhsflektavlrvpvppnvaf 120
·eldrtradllisedvqrrfvqevvqsqqvavgrqledfrsk 160
rlmgmtpweqelaqleawvgrdrasyearerhvaerllmh 200
             210       ·220       230       240
    ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

leemqhtistdeeksaavvnaiglymrhlgvrtksgdkks 240
grnffrkkvmgnrrsddppktkkglssildaarwnrgepq 280
vpdfrhlkaevdaekpgatdrkggvgmpsrdrnigapgqd 320
tpgvslhplsldspdrepgadapleigdsspqgpmsles l 360
appestdegaetespepgdegepgrsglelepeeppgwre 400
             410       420       430       440
    ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

lvppdtlhslpksqvkrqevisellvteaahvrmlrvlhd 440
lffqpmaeclffpleelqnifpsldelievhslfldrlmk 480
rrqesgylieeigdvllarfdgaegswfqkissrfcsrqs 520
faleqlkakqrkdprfcafvqeaesrprcrrlqlkdmipt 560
emqritkyplllqsigqnteepterekvelaaeccreilh 600
             610       620       630       640
    ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴_____

hvnqavrdmedllrlkdyqrrldlshlrqssdpmlsefkn 640
lditkkklvhegpltwrvtkdkavevhvllldd llll lqr 680
qderlllkshsrtltptpdgktmlrpvlrltsamtrevat 720
dhkafyvlftwdqeaqiyelvaqtvserknwcalitetag 760
slkvpapasrpkprprpsstrepllsssengnggretspa 800
             810       820       830       840
    ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴_____

darterilsdllpfcrpgpegqlaatalrkvlslkqllfp 840
aeedngagpprdgdgvpgggplspartqeiqenllsleet 880
mkqleeleeefcrlrpllsqlggnsvpqpgct 912
```

FIGURE 10

```
              10        20        30        40
        |    |    |    |    |    |    |    |    |    |

gggcgccccgccggtcacttccgcgcggacaccagccttg  40
cagagcccaggggagatggaagacttcgcccgaggggcggc  80
ctccccaggcccctcccggcctggcctggttcccgtcagc  120
atcatcggggctgaggatgaggattttgagaacgagctgg  160
agacaaactcagaagagcaaaacagccagttccagagcct  200

             210       220       230       240
        |    |    |    |    |    |    |    |    |    |

ggagcaggtgaagcggcgcccagcccacctcatggccctc  240
ctgcagcacgtggccctgcagtttgagccaggacccctgc  280
tttgctgtctgcatgccgacatgctgggctcactgggccc  320
caaggaggccaagaaggccttcctggacttctaccacagc  360
ttcctggagaagacagcggttctccgggtgccggtccctc  400

             410       420       430       440
        |    |    |    |    |    |    |    |    |    |

ccaacgtcgcctttgaacttgaccgcactagggctgacct  440
catctccgaggatgtccagcggcggttcgtgcaggaggtg  480
gtgcaaagccagcaggtagccgtgggccggcagctggagg  520
acttccgttccaagcggctcatgggcatgacgccctggga  560
gcaggagctggcccagctggaggcttggggttgggcgggac  600

             610       620       630       640
        |    |    |    |    |    |    |    |    |    |

cgagccagctacgaggcccgggagcggcacgtygcggagc  640
ggctgctcatgcacctggaggagatgcaacataccatctc  680
taccgacgaagaaaagagtgctgccgtggtcaabgccatt  720
gggctgtacatgcgccaccttggggtgcggaccaagagtg  760
gagacaagaagtcggggaggaacttcttccggaaaaaggt  800

             810       820       830       840
        |    |    |    |    |    |    |    |    |    |

gatggggaaccggcggtcggacgaccctcccaagaccaag  840
aaggggctgagcagcatcctggatgccgcccgctggaacc  880
ggggagagccccaggttccagattttcgacacctcaaagc  920
agaggttgatgccgagaagccaggtgctacagaccggaag  960
ggaggcgtggggatgccctctcgggaccggaatatcgggg  1000

            1010      1020      1030      1040
        |    |    |    |    |    |    |    |    |    |

ctcctgggcaggacacccctggagtctctctgcaccctct  1040
gtccctggacagcccagaccgggaaccaggtgctgacgcc  1080
ccctggagctggggggactcatccccgcagggcccaatga  1120
gcctggagtccttggcgcccccagagagtaccgacgaggg  1160
ggccgaaaccgagagccccgagcctggagatgaggggggag  1200

            1210      1220      1230      1240
        |    |    |    |    |    |    |    |    |    |

ccggggcggtcgggactggagcttgaaccagaagagcctc  1240
ccggctggcgggaactcgtcccccccagacaccctgcacag  1280
cctgcccaagagccaggtgaagcggcaggaggtcatcagc  1320
gagctgctggtgacagaggcggcccacgtgcgcatgctgc  1360
gggtgctgcacgacctcttcttccagcccatggcagaatg  1400
```

FIGURE 11

SUBSTITUTE SHEET (RULE 26)

```
              1410        1420        1430        1440
     |‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|
     cctgttcttccccttggaggagctgcagaacatcttcccc   1440
     agcctggacgagctcatcgaggtgcattccctgttcctcg   1480
     atcgcctgatgaagcggaggcaggagagtggctacctcat   1520
     cgaggagatcggagacgtgctgctggcccggtttgatggt   1560
     gctgagggctcctggttccagaaaatctcctcccgcttct   1600

              1610        1620        1630        1640
     |‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|
     gcagccgccagtcatttgccttagagcagctcaaagccaa   1640
     gcaacgcaaggaccctcggttctgtgccttcgtgcaggaa   1680
     gctgagagccgcccgcggtgccgccgcctgcagctgaagg   1720
     acatgatccccacggagatgcagcggctgaccaagtaccc   1760
     cctgctcctgcagagcatcgggcagaacacagaagagccc   1800

              1810        1820        1830        1840
     |‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|
     acagaacgggagaaagtggagctggcagccgagtgctgcc   1840
     gggaaattctacaccacgtcaaccaagccgtgcgtgacat   1880
     ggaggacctgctgaggctcaaggactatcagcggcgcctg   1920
     gacttgtcccaccttcggcagagcagcgaccctatgctga   1960
     gcgagttcaagaacctggacatcaccaagaagaaattggt   2000

              2010        2020        2030        2040
     |‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|
     ccacgagggcccactgacgtggcgggtgactaaggacaag   2040
     gcagtggaggtgcatgtgctgctgctggacgacctgctgc   2080
     tgctgctccagcgccaggacgagcggctgctgctcaagtc   2120
     ccatagccggacactgacgcccacgcccgatggcaagacc   2160
     atgctgcggcccgtgctgcggctcacctccgccatgaccc   2200

              2210        2220        2230        2240
     |‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|
     gcgaggtggccaccgatcacaaagccttctacgtccttttt   2240
     tacctgggaccaggaggcccagatatacgagctggtggca   2280
     cagactgtgtcggagcggaaaaactggtgtgctctcatca   2320
     ctgagactgccggatccctgaaagtccctgcccctgcctc   2360
     tcgccctaagccccggcccaggccgagcagcacccgagaa   2400

              2410        2420        2430        2440
     |‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|
     cccctcctcagcagctctgagaacgggaatggtggccgag   2440
     agacgtctccagctgatgcccggaccgagagaatcctcag   2480
     tgacctcctgcccttctgcagaccaggccccgagggccag   2520
     ctcgctgccacggcccttcggaaagtgctgtccctgaagc   2560
     agcttctgtttccggcggaggaagacaatggggcggggcc   2600

              2610        2620        2630        2640
     |‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|‥‥‥|
     tcctcgagatggggatggggtcccaggggggcgggcccctg   2640
     agcccagcacggacccaggaaatccaggagaacctgctca   2680
     gcttggaggagaccatgaagcagctggaggagttggagga   2720
     ggaattttgccgcctgagacccctcctgtctcagcttggg   2760
     gggaactctgtcccccagcctggctgcacttgaggttccc   2800
```

FIGURE 11 cont.

SUBSTITUTE SHEET (RULE 26)

```
              2810         2820         2830         2840
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      gcccaggaaggccttttgcaagaaggagaggaatgggggga 2840
      gaggacgtgagggaccacccccacccacacagctgccgca 2880
      gcatctcacaccccgagggcctgaggagagggagctgtgg 2920
      gccacgcctgggaggggcccagctggggttactgcccccg 2960
      catgagcctcggccatctctccctcctgccctctgcttgg 3000

              3010         3020         3030         3040
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      gggactcagggctccattctggagggcaccacggtgaccc 3040
      gggccatctcagtattgcctgtgggggccacccctccacc 3080
      cccacccccaagtgccttcgctctgttttttataccctgaa 3120
      ttggagggtttatttttttaatatatatattat 3150
```

FIGURE 11 cont.

```
                10          20          30          40
       ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴

msvrlpqsidrlssisslgdsaperkspshhrqpsdaset  40
tglvqrcviiqkdqhgfgftvsgdrivlvqsvrpggaamk  80
agvkegdriikvngtmvtnsshlevvkliksgayvaltll  120
gsspssmgisglqqdpspagapritsvipsppppplppp  160
qritgpkplqdpevqkhatqilrnmlrqeekelqdilply  200

               210         220         230         240
       ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴

gdtsqrpsegrlsldsqegdsgldsgterfpslseslmnr  240
nsvlsdpgldsprtspvimarvaqhhrrqgsdaavpstgd  280
qgvdqspkpliigpeedydpgyfnnesdiifqdleklksr  320
pahlgvflryifsqadpspllfylcaevyqqaspkdsrsl  360
gkdiwnifleknaplrvkipemlqaeidsrlrnsedargv  400

               410         420         430         440
       ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴

lceaqeaampeiqeqihdyrtkrtlglgslygendiidid  440
gdplrerqvaekqlaalgdilskyeedrsapmdfalntym  480
shagirlrearpsntaekaqsapdkdkwlpffpktkkssn  520
skkekdaledkkrnpiikyigkpksssqstfhiplspvev  560
kpgnvrniiqhfennqaydapepgtqrlstgsfpedlles  600

               610         620         630         640
       ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴

dssrseirlgrseslkgreemkrsrkaenvprsrsdvdmd  640
aaaeatrlhqsassstsslstrslenptppftpkmgrrsi  680
espslgfctdtllphlleddlgqlsdlepepdaqnwqhtv  720
gkdvvagltqreidrqevinelfvteashlrtlrvldlif  760
yqrmkkenlmpreelarlfpnlpelieihnswceamkklr  800

               810         820         830         840
       ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴

eegpiikeisdlmlarfdgpareelqqvaaqfcsyqsial  840
eliktkqrkesrfqlfmqeaeshpqcrrlqlrdliisemq  880
rltkypillesiikhteggtseheklcrardqcreiikyv  920
neavkqtenrhrlegyqkrldatalerasnplaaefksld  960
lttrkmihegpltwriskdktldlhvllledllvllqkqd  1000

               1010        1020        1030        1040
       ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴

eklllkchsktavgssdskqtfspvlklnavlirsvatdk  1040
raffiictsklgppqiyelvaltssdkntwmelleeavrn  1080
atrhpgaapmpvhppppgprepaqqgptpsrvelddsdvf  1120
hgepepeelpggtgsqqrvqgkhqvlledpeqegsaeeee  1160
lgvlpcpstsldgenrgirtrnpihlafpgplfmeglads  1200

               1210        1220        1230        1240
       ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴  ┴

aledvenlrhliilwsllpghtmetqaaqepeddltptpsv  1240
isvtshpwdpgspgqappggegdntqlaglegerpeqedm  1280
glcslehlpprtrnsgiwespeldrnlaedassteaaggy  1320
kvvrkaevagskvvpalpesgqsepgppeveggtkatgnc  1360
fyvsmpsgppdsstdhseapmsppqpdslpagqtepqpql  1400
```

FIGURE 12

```
                 1410        1420        1430        1440
        |    |    |    |    |    |    |    |    |    |    |
        qggnddprrpsrspps lal rdvgmi fhtieq lt lk lnr lk  1440
        dme lahre l lks l ggessggttpvgsfhteaarwtdgs ls  1480
        ppakep lasdsrnshe l gpcpedgsdap ledstadaaasp     1520
        gp  1522
```

FIGURE 12 cont.

16/34

```
              10        20        30        40
    |....|....|....|....|....|....|....|....|
aattggctcatttaagaatttcaaaacatttaatgtaaaa 40
gcttttttttttttaaggaagtccataaattttggttccc 80
agggttgcactggacttggaaggagtgctgttgtgtacat 120
actattgtatggtttattttattattttactgtacaaatc 160
agccgaaagaattttttccaagtgccatttcggatttatta 200
             210       220       230       240
    |....|....|....|....|....|....|....|....|
atccttttttttttcctttcctcaaagatatttgctgttg 240
tcatattaagcattggagactagaaaattactttcccct 280
ttgagctagagggtctcttgccaacagaaggacagctgag 320
aaagctggatttaaaggatggttttatctgtactttgcag 360
ttaacagtgatattttgaaggcacattttttctgtgattca 400
             410       420       430       440
    |....|....|....|....|....|....|....|....|
tttttttttggccatagtgctaaccttgaagagattcgtg 440
gctgggtttttggtttctgagaaggtcgtagttttttcctc 480
ttttccttttttttttttcttttttcttttctttttcttttt 520
ttttaaagcggggggagggaagaggggctgagaaggaaa 560
tcatgttcactggtagaagtagagtggagcatcagttacc 600
             610       620       630       640
    |....|....|....|....|....|....|....|....|
agggtcctgagagctggaggcgaaaggattctatcttcaa 640
gttgggaggcccctcctctccaccttgctcaaaaattgcaag 680
cgattcaatcctgatcaagacaccaaagctacaggattct 720
ggaaccgtggagacaccgagaaaccatgagtgtaaggtta 760
ccccagagtatagacaggttaagtagcctgtcttctctgg 800
             810       820       830       840
    |....|....|....|....|....|....|....|....|
gagattctgcaccagagcgcaagtccccttcccaccatcg 840
ccagccttcggatgcctctgagacaacaggtctcgttcaa 880
cgctgtgtcattatccaaaaggaccagcatggcttcggct 920
tcacagtcagtggggatcgcattgttctggtgcagtctgt 960
gcggcctggaggtgcagccatgaaggccggtgtgaaagag 1000
             1010      1020      1030      1040
    |....|....|....|....|....|....|....|....|
ggcgaccggatcatcaaagtcaacggcaccatggtgacca 1040
atagctcacacctggaagtggtaaagctgatcaaatctgg 1080
cgcctatgtcgcactcaccctcctgggctcttcaccttca 1120
tccatgggcatctctgggctccagcaggacccatccccag 1160
caggagctccccgaatcacgtcagtgatcccctcaccacc 1200
             1210      1220      1230      1240
    |....|....|....|....|....|....|....|....|
acctcctcccacctctaccacctccacaacgcatcacagga 1240
cccaaacctctgcaggatcccgaagttcaaaaacatgcca 1280
cccagatcctcaggaatatgctgaggcaggaagaaaaaga 1320
attacaggacatacttccactatatggtgacaccagccag 1360
agaccatcagaaggccggctctctctggattcccaggagg 1400
```

FIGURE 13.

```
                1410       1420       1430       1440
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        gggacagtggcttggactctgggacagaacgctttccttc 1440
        cctcagtgagtcattgatgaatcggaactcggtactgtca 1480
        gaccctgggctagacagtcctcgaacctcccctgtgatca 1520
        tggccagggtggcccagcaccacaggcggcagggctcgga 1560
        tgcagcagtcccctcaaccggtgaccagggtgtagatcaa 1600

                1610       1620       1630       1640
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        agcccaaagcctttaattattggcccagaggaagactatg 1640
        acccgggttatttcaacaacgagagcgacatcatattcca 1680
        ggatctggagaaactgaagtctcggccagctcacctgggg 1720
        gtttttctacgttacatcttctctcaggcggaccccagtc 1760
        cactgcttttttacctgtgtgcagaagtttatcagcaggc 1800

                1810       1820       1830       1840
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        aagccccaaggattcccgaagcttggggaaagacatctgg 1840
        aatattttcctggagaaaaatgcgcctctgagagtgaaga 1880
        tccctgagatgctacaggctgaaattgactcgcgcctgcg 1920
        gaacagcgaagatgcccgtggtgttctctgtgaagctcaa 1960
        gaggcagccatgcctgagatccaagagcagatccacgact 2000

                2010       2020       2030       2040
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        acagaacgaagcgcacactggggctgggcagcctgtatgg 2040
        tgaaaatgacctgctggacctggatggggaccctctccga 2080
        gagcgccaagtggctgagaagcagctggctgcccttggag 2120
        atattttgtccaagtatgaggaagacaggagcgcccccat 2160
        ggacttcgccctcaatacctacatgagccatgctgggatc 2200

                2210       2220       2230       2240
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        cgtcttcgagaggcacgaccttccaacacagctgaaaagg 2240
        cccagtctgctcctgacaaggacaagtggctaccgttctt 2280
        ccctaagaccaagaagagcagcaattccaagaaagaaaag 2320
        gatgccttggaggacaagaagcgaaaccctatcctcaaat 2360
        acattgggaagcccaaaagctcttctcaaagcacatttca 2400

                2410       2420       2430       2440
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        tattcccttgtcccctgtggaagtcaaaccaggcaatgtg 2440
        aggaacatcattcagcactttgagaacaaccagcagtatg 2480
        atgccccagaacctgggacacaacgactctcgaccggaag 2520
        cttttcctgaggacctgctggagagtgacagttcacgctca 2560
        gagattcgcctgggccgctctgaaagcctcaagggccggg 2600

                2610       2620       2630       2640
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        aagagatgaaacggtctcgaaaggcagagaacgtgccccg 2640
        ctctcgcagtgatgttgacatggatgctgctgcggaggct 2680
        actcgcctgcaccagtcagcctcgtcctctacctccagcc 2720
        tctccaccaggtctcttgagaacccaaccccctccattcac 2760
        tcccaaaatgggccgcaggagcattgagtcccccagtttg 2800
```

FIGURE 13 cont.

```
            2810        2820        2830        2840
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      gggttctgcacagataccctccttccccacctcctagagg 2840
      atgatctgggccagctgtctgacctggagccagagccaga 2880
      tgcccaaaattggcagcatacagtgggcaaggatgtggtg 2920
      gctgggctaacccagcgggagattgaccggcaagaggtca 2960
      tcaatgagctgtttgtgactgaagcttcccacctgcgcac 3000

            3010        3020        3030        3040
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      actccgggtcctggacctgatcttctaccagcgaatgaag 3040
      aaggagaacctgatgccccgggaggagctggcccggctct 3080
      tcccgaacctgcctgaactcatagagattcacaattcctg 3120
      gtgtgaagccatgaagaagctccgggaggaaggccccatc 3160
      atcaaagagatcagtgacctcatgctggcccggtttgatg 3200

            3210        3220        3230        3240
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      gccctgcccgagaggaactccagcaagtggctgcacagtt 3240
      ctgttcctatcagtcaatagccctagagctaatcaagacc 3280
      aagcaacgcaaggagagtcgattccagctcttcatgcagg 3320
      aggctgagagccaccctcagtgtcggcggctgcagctgag 3360
      agacctcatcatctctgagatgcagcggctcaccaagtac 3400

            3410        3420        3430        3440
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      ccgctgctgctggagagcatcatcaagcacacagaggggtg 3440
      gcacctctgagcatgagaagctgtgccgggcccgggacca 3480
      gtgccgggagattctcaagtatgtgaatgaagcggtaaaa 3520
      caaacagagaaccgccaccgtttagagggctaccagaaac 3560
      gcctggatgccaccgccctggagagggccagcaaccccct 3600

            3610        3620        3630        3640
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      ggcagcagagttcaagagcctggatcttacaaccagaaaa 3640
      atgatccatgagggacccctgacctggaggatcagcaagg 3680
      ataagaccttggacctccacgtgctgctgctggaggacct 3720
      cctagtgctgctacagaaacaggatgagaagctattgctg 3760
      aagtgccacagcaagactgctgtgggctcctcagacagca 3800

            3810        3820        3830        3840
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      agcagaccttcagccccgtgctcaagctcaatgctgtgct 3840
      catccgctctgtggccacagataaacgggccttcttcatc 3880
      atctgcacctccaagctgggcccacccagatctatgagc 3920
      tggttgcattgacgtcatcagacaagaacacatggatgga 3960
      gctcttagaagaggccgtgcggaatgccaccaggcacccc 4000

            4010        4020        4030        4040
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

      ggagctgcccccaatgcccgtccatcctccacccccaggtc 4040
      cccgggagccagccagcagggcccacacccagcagggt 4080
      agaactggatgactcagacgtgttccatggtgaacctgaa 4120
      cctgaggagctgcctggaggcactgggtcccagcagaggg 4160
      tccaagggaagcaccaggtcctgctagaggaccctgagca 4200
```

FIGURE 13 cont.

```
              4210          4220          4230          4240
     |    |    |    |    |    |    |    |    |    |    |    |
     ggagggcagtgcagaggaagaggaactgggtgtcctgcct  4240
     tgcccttccacatccctggatggagagaacaggggcatca  4280
     ggacaaggaaccccatccacttggccttcccaggccctct  4320
     gttcatggaagggctcgctgactccgctctggaagatgtg  4360
     gagaacctgcgacatctgatcctgtggagcctgctgccag  4400
              4410          4420          4430          4440
     |    |    |    |    |    |    |    |    |    |    |    |
     gtcacaccatggaaactcaggctgcccaggagcccgagga  4440
     cgacctgacacccacaccttctgtcatcagcgtcacctct  4480
     cacccctgggacccaggctccccagggcaagcacccctg   4520
     ggggtgaaggggacaacacccagcttgcagggctggaggg  4560
     ggaacggccagagcaggaagacatgggtctctgttctctg  4600
              4610          4620          4630          4640
     |    |    |    |    |    |    |    |    |    |    |    |
     gaacacctaccccccaaggaccagaaattctgggatatggg  4640
     agtctccagaactggacaggaatctggctgaagatgcttc  4680
     aagcacagaggcagcaggaggttacaaagttgtgagaaaa  4720
     gctgaggtggcaggcagcaaggttgtccctgcactaccag  4760
     agagtggccagtcagagcctgggccacctgaagtggaagg  4800
              4810          4820          4830          4840
     |    |    |    |    |    |    |    |    |    |    |    |
     cggaacaaaggctacggggauctgcttttatgtcagcatg  4840
     ccatcaggaccccggactcaagcaccgaccactcagagg   4880
     cacccatgagcccccctcagcctgacagcctccctgcagg  4920
     gcagacagagcctcagcctcagctgcagggaggcaacgat  4960
     gatccaagacgccccagccgctctcctccaagcctggccc  5000
              5010          5020          5030          5040
     |    |    |    |    |    |    |    |    |    |    |    |
     tcagggacgtgggcatgatcttccataccattgagcagct  5040
     cactctcaagctcaacaggctcaaggatatggagctggcc  5080
     cacagagagctgctcaagtcccttggggagagtcatctg   5120
     gtggcaccacgcctgtgggcagtttccacacagaagcagc  5160
     tagatggacagatggctccctctcacctcccgctaaggag  5200
              5210          5220          5230          5240
     |    |    |    |    |    |    |    |    |    |    |    |
     cccctagcttctgactccaggaacagccatgaactggggc  5240
     cctgccctgaggatggctctgacgcccccctggaagacag  5280
     cacagcagacgcagccgcgtcaccaggaccataaccgtac  5320
     aaaccaccaaatcctctgcgtccccactcctccttcaggg  5360
     actggcctgagaccggggcacagggtaggggggatcccaa  5400
              5410          5420          5430          5440
     |    |    |    |    |    |    |    |    |    |    |    |
     cactcctccctgtggaggaggcagttagggaaactaggat  5440
     ccagccaaggcccggggggagacccgcatgttgcttggtc  5480
     tgctcaagtcggagtcaggtttcagtgtcttttccctccc  5520
     ttagcccaaccctccaaggcctcatgtctcctaagcatgc  5560
     tgactgcatccgaaaggcccccactcaccatggtctgccc  5600
```

FIGURE 13 cont.

SUBSTITUTE SHEET (RULE 26)

```
            5610        5620        5630        5640
    |...|...|...|...|...|...|...|...|...|...|...|_____

    tcaccccacatatgtgtgtacacgcgcacgcctgtatgtg 5640
    cgctgccctcagacatgcaagtgaaaggaggaggcttctg 5680
    tgtaaatgcactttcttcctcccctctttctccataagac 5720
    cccaggcagaggtgggtgcctcccctcccctctttgtcac 5760
    tttggtttcctataaatatgtatgtatcgt 5790
```

FIGURE 13 cont.

```
                10        20        30        40
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        mgevaggaapgpprsglvsiiigaededfeneleansedq 40
        nsqfqsleqvkrrpahlmallqhvalqfepgpllcclhad 80
        mlsslgpkeakkafldfyhsflektavlrvpvppsvafel 120
        drtrpdlisedvqrrfiqevvqsqqaavsrqledfrskrl 160
        mgmtpweqelsllepwigkdrgnyearerhvaerllshle 200
                210       220       230       240
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        etqhtistdeeksaavvtaislymrhlgvrtksgdkksgr 240
        nffrkkvmgnrrsdeppktkkglssildparwnrgepsap 280
        dcrhlkveadekpgpadrkgglgmssrdrtvgtpgqdnpg 320
        vsihplstdsvdsrepgvdtpqepgdtppqgptsleplap 360
        pestedngetespepgddgepgrsglelepeeppgwrelv 400
                410       420       430       440
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        ppdtllslpksqvkrqevisellvteaahvrmlrvlhdlf 440
        yqpmadggffpldelqnifpsldelievhslfldrlmkrr 480
        qesgylieeigdvllarfdgaegswfqkissrfcsrqsfa 520
        leqlkakqrkeprfcafvqeaesrprcrrlqlkdmiptem 560
        qrltkyplllqsigqnteestergkvelaaeccreilhhv 600
                610       620       630       640
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        nqavrdmedllrlkdyqrrldlthlrqssdpmlsefknld 640
        itkkklvhegpltwrvtkdkaievhvllldllllllqrqd 680
        erlllkshsrtltptpdgktmlrpvlrltsamtrevatdh 720
        kafyviftwdqeaqiyelvaqtsserknwcnlitetagsl 760
        kvpapasrlkprpspssirepllsssengtggaemapada 800
                810       820       830       840
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        rterllndllpfcrpgpegqlaatalqkvislkqillste 840
        edsgagpprdgdgvpggrapgpvhtqeieenllslevair 880
        qleeleeefcrirpllsqlggtlspnlaapersaqtgls 919
```

FIGURE 14

```
                10          20          30          40
        |........|........|........|........|........|

        cgggacaccgggcccgggatccccgagcccgacctcgggc  40
        gcctcgccggtcacctccccgcgcggacaccagcccggcg  80
        gagtcctggagatggggagaagtcgccggaggggcggcccc 120
        agggcctccccggtctggcctggtgtccatcatcatcggg  160
        gcggaggatgaggattttgagaacgagctggaggcgaact  200

               210         220         230         240
        |........|........|........|........|........|

        cagaagatcaaaacagccagttccagagcctagagcaagt  240
        gaagcgccgccctgcccacctcatggccctcctgcagcat  280
        gtggccctgcagttcgagccaggaccactgctctgctgcc  320
        tgcatgcagacatgctgagctctctgggccccaaagaagc  360
        caagaaggccttccttgacttctatcacagtttcctggag  400

               410         420         430         440
        |........|........|........|........|........|

        aagactgcggttctacgggtgccggtccctcccagtgtcg  440
        cttttgaacttgatcgtactcgacctgatctgatctctga  480
        ggatgtccagaggcggttcatacaagaggtggtgcagagc  520
        cagcaggcagccgtgagccgtcagctagaggacttccgct  560
        ccaagcggctcatgggcatgacgccctgggagcaggaact  600

               610         620         630         640
        |........|........|........|........|........|

        gagcctgctggagccctggatgggaaagaccgaggcaac   640
        tatgaggcccggggagcggcatgttgcggagcggctgctgt 680
        cccacctggaggagacccagcataccatctctacagatga  720
        agagaaaagtgctgctgtggtcactgccatcagcctgtat  760
        atgcgccaccttggagtccggaccaagagtggggacaaga  800

               810         820         830         840
        |........|........|........|........|........|

        agtcgggaaggaacttcttccggaaaaaggtgatggggaa  840
        tcggaggtcagacgaacccccaaagacaaagaaagggctg  880
        agcagtatcctagatcctgcacgttggaaccgggggagagc 920
        catccgctccagattgtcgacatctaaaggtcgaggctga  960
        tgagaagccaggccctgcagaccggaagggaggcctgggt 1000

              1010        1020        1030        1040
        |........|........|........|........|........|

        atgtcttctcgggacaggactgttgggactcctggacagg 1040
        acaacccaggagtctccctgcaccctctgtctacagacag 1080
        cgtcgactcccgggaaccaggcgtggataccccgcaggag 1120
        ccaggggatacaccccacagggccctaccagcctggagc  1160
        ccctggcgccccagagagcacagaggacaatggcgagac  1200

              1210        1220        1230        1240
        |........|........|........|........|........|

        tgagagccctgagcccggagatgatggggagccaggacgg 1240
        tcaggcctggaactggaaccagaagaacctcctgggtgga 1280
        gggaactcgtgccccagacaccctgctcagtctgcccaa  1320
        gagccaagtgaagcggcaagaggtcatcagcgagctgctc 1360
        gtgactgaggcagctcacgtgcgcatgctacgggtactgc 1400
```

FIGURE 15

```
                   1410      1420      1430      1440
         ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
         atgacctcttctaccagcccatggcggatggaggcttctt  1440
         ccctctggacgagctgcagaacatcttcccgagcctggat  1480
         gagctcatcgaggtgcactccctgttcctcgatcgcttga  1520
         tgaagcggagacaagagagtggctacctcattgaggagat  1560
         cggcgatgtgctactggcccggttcgatggtgctgagggc  1600
                   1610      1620      1630      1640
         ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
         tcatggttccagaagatctcctcccgcttctgcagccgcc  1640
         agtcgttcgctctagagcagctcaaagccaagcagcgcaa  1680
         ggagcctcggttctgtgcctttgtgcaggaagctgagagc  1720
         cgcccgagatgccggcgcctacagttaaaggacatgatcc  1760
         ccactgagatgcagcgactgaccaagtacccactgctgct  1800
                   1810      1820      1830      1840
         ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
         acagagcatcgggcagaacacagaggagtctacagaacga  1840
         gggaaagtggagcttgcagctgagtgctgccgggagattc  1880
         tgcaccatgtcaatcaagccgtccgtgacatggaggacct  1920
         gctgcggctcaaggattaccagcggcgcctggacttgact  1960
         cacctacggcagagcagtgaccctatgctgagcgagttca  2000
                   2010      2020      2030      2040
         ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
         agaacctggacatcactaagaagaagttggtrcatgaagg  2040
         ccccctcacgtggcgagtgaccaaagacaaagctatagaa  2080
         gtgcacgtgctcttgctggacgacctgctgctgctgctcc  2120
         agcgccaggacgagaggctgctgctcaagtcccacagccg  2160
         gacgctgacacctaccccgatggcaagaccatgctgcgg  2200
                   2210      2220      2230      2240
         ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
         ccggtgctccggctcacctctgccatgacccgagaggtgg  2240
         ccactgatcacaaagctttctacgtcatttttacctggga  2280
         ccaggaggcccagatatatgagctggtggcacagacatct  2320
         tcggaacgcaaaaactggtgtaacctcatcactgagactg  2360
         ctggatccctgaaggtccctgccctgcctcccgcctcaa  2400
                   2410      2420      2430      2440
         ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
         accccggcccagcccaagcagcatccgagaacccctgctc  2440
         agcagctctgagaatggcactggaggcgcagagatggctc  2480
         cagctgatgccaggacagagcggctcctcaatgacctcct  2520
         gcccttctgcagaccaggcccagagggccagcttgctgcc  2560
         acagcccttcagaaagtactgtccctgaagcagatcctgc  2600
                   2610      2620      2630      2640
         ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
         taagcactgaggaagacagtggagcggggcctccccgcga  2640
         tggggatggggtgcctggtggtagggcccccggcccagtg  2680
         cacacccaggagattgaggaaaacttgcttagcttagagg  2720
         tggccatcagacaactggaggagttggaagaggaattttg  2760
         tcgcctaagacccctcctgtcccagcttggggggactctg  2800
```

FIGURE 15 cont.

```
              2810        2820        2830        2840
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
        tcccccaacctggctgcacctgaacgctctgctcagacag 2840
        gcctttcatgaagagaagagtgggggagaggatgcagggg 2880
        ggcccaccccccacccccacagctgccacagcatctcacac 2920
        cccagaggcctggagcagagggagattggctgaacttgat 2960
        agcagacctgctcgggtccctgcctcccactctggcctct 3000
              3010        3020        3030        3040
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
        gcctttctccctcctgccttctgcttgggggaactcagggc 3040
        ttcactcccagagcaccccttaatccccacttttcaggc 3080
        cgtctcagtattgcctatgggggggtcaccctccttcccct 3120
        accccaagtgcctttgctatgtttttatatcctggactgg 3160
        aggtttatttttaatatatattatctaaggaaaaaaaaa 3200
```

FIGURE 15 cont.

```
            10          20          30          40
  ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

mddpsikkrlldlytdeheydevqeipeessiqppetsts  40
htstngsshsgpgtatgpgatsagpsagapqspvivvdsv  80
pelpapkqksvknskskqkqkqlankskiprspslassls 120
slasslsghrdrdkdrdkdrenqnavppqtpplppsykqn 160
qmngdstaaagggvsapatpttannnnashnngsimgggv 200

           210         220         230         240
  ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

qlnqsdnsnpvlqapgerssinltplsrdlsgghtqestt 240
pattpstpslalpknfqyltltvrkdsngygmkvsgdnpv 280
fvesvkpggaaeiaglvagdmillrvnghevrlekhptvvg 320
likasttvelavkrsqkltrpssvsvvtpstpilsgrdrt 360
asitgpqpvdsikrremetykiqtlqkmleqeklnlerlk 400

           410         420         430         440
  ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

sdqnnpsyklseanirklreqlhqvgaedaptvklqaaag 440
nkntalltpnqiqhlsasathsnqqfhhlhhhhnlhnnny 480
ppqqqpastspaflsllprslsslslgtrknktekdltts 520
spfglttdflqqqrmshqaesmsqsmhqhtstptsqqffh 560
phqqqhrfketgptskgknkfllisrsllieedvppplpqrn 600

           610         620         630         640
  ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

pprqlnldlkngnaspggshlvapvsdldratspqlnrsq 640
qqqlprstdnspsnakskrskiktkalsdpkmstqmllqm 680
esasaagaaggsievdggppplpprlpgmmtedmsrgscq 720
nlaqpnsvgtafnyplvstttavqndnlniafplsqrpni 760
vqqlqqyqqqqhqmsggqatgalgqtpnlgknkhrrvgs 800

           810         820         830         840
  ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

spdnmhprhpdritkttsgsweivekdgessppgtpplpy 840
lssshmtvledpnennrgaaaagpgvfieshqftpmagas 880
spipislhsshmhaaqsndtqkeiismedensdldepfid 920
engpfnnltrlleaenvtflaiflnyvisnsdpaplLfyl 960
itelykegtskdmrkwayeihstflvpraplswyrqdesl 1000

           1010        1020        1030        1040
  ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

arevdnvlqleydkveilrtvflrsrkrakdliseqiref 1040
qqkrtaglgtiygptddklaeaktdklreqiidkylmpnl 1080
haliedengsppedvrkvalcsalstviyrifntrpppss 1120
ivervhhfvsrdksfksrimgknrkmnvrghplvlrqyye 1160
vthcnhcqtiiwgvspqgyhctdcklnihrqcskvvdesc 1200

           1210        1220        1230        1240
  ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

pgplpqakrlahndkiskfmgkirprtsdvignekrsrqd 1240
eeldveltpdrgqasivrqpsdrrpdanisirsngntscn 1280
tsglnttdlqssfhgscandsinpgggagcnmdlstsvas 1320
ttpstsgsvaaglsafaelnaldtvdkearrerysqhpeh 1360
ksapvsvnrsesykerlsnkrnrnsrrktsdpslssrpnd 1400
```

FIGURE 16

```
              1410        1420        1430        1440
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        eqldlglsnatyvgssnsslssaggsespstsmehfaapg 1440
        aaggvqvppmglnqnqhphlliqqhaqqycqqdsfqagla 1480
        gaagssaasnssfwnaghplpvarwtlesededdvneadw 1520
        ssmvaaevlaaltdaekkrqeiineiyqternhvrtlkll 1560
        drlfflplyesgllsqdhlllifppallslreihgafeqs 1600

              1610        1620        1630        1640
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        lkqrriehnhvvntigdlladmfdgqsgvvlcefaaqfca 1640
        rqqialealkekrnkdemlqkllkkseshkacrrlelkdl 1680
        lptvlqrltkypllfenlykvtvrllpentteaeaiqrav 1720
        esskrillvevnqavktaedahklqniqrkldrssydkeef 1760
        kkldltqhhlihdgnltikknpsvqlhgllfenmivlltk 1800

              1810        1820        1830        1840
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        qddkyylknlhtplsitnkpvspimsidadtlirqeaadk 1840
        nsfflikmktsqmlelrapsssecktwfkhfsdvaarqsk 1880
        nrsknassnhdtsisdpalaaiphsntkeslelstdtvqp 1920
        laatatltttplapmlpiatvtpapatnnsnvssltgvql 1960
        rnpqrdatasesdadyvntpkprssqnevnrtmsirstge 2000

              2010        2020        2030        2040
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        piqkysangteandvtlrhsqstresvrpgstgeernsty 2040
        gmvggnskrdsasivcsnnsnntrtllmqsplvdptaiqv 2080
        sispahtaepvltpgeklrrldasirndllekqkiicdif 2120
        rlpvehydqivdiammpeapkdsadialaaydqiqtltkm 2160
        lneymhvtpeqevsavstavcghchekeklrkkvapsssf 2200

              2210        2220        2230        2240
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        sssppplpppnrqhaqaqaqippsrlmpklqtldldevai 2240
        hedddgyceidelrlpaipskpherpttplapfntepkts 2280
        qsvidaskrqstdavpeglleqeplegdktetkgednevk 2320
        tvpsdklsescneerqcveaditkevadpttskneaaasv 2360
        delpsqsreiktaenasksvadkkedneetieegvastvd 2400

              2410        2420        2430        2440
        └┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┘

        sstqtsptespketdkltggssstcgpnriqhasvlepsv 2440
        pchalssivtilneqismllpkinerdmererlrkenqhl 2480
        rellsalhdrqrvdevketpfdlkklmhaedvefdddida 2520
        isnssltptptpiptaspsasgqvetaeamritstedee 2559
```

FIGURE 16 cont.

```
                10        20        30        40
     |....|....|....|....|....|....|....|....|
     aagaacgtcgtgtgcgtgttgatggcgcgataagtgaaaa 40
     aaataaaaacatttttcgtgtctttatggaatacaatacg 80
     agtgcgttttaagtttaagtgaaaagtagcagtgcaaaat 120
     atatatacacatatcaaaacccattggaatagtgcaacca 160
     aaaaaacataaaacagaactaaagcaacgaaaatcgaggc 200
                210       220       230       240
     |....|....|....|....|....|....|....|....|
     aattaaatcacttctggattcatttcggcccgttttttact 240
     agctaatttaaaaatagctaccgaaagcattccgcaacca 280
     cttctggtgactggccccacaccaatgcgaacaatactcc 320
     acgggatcctccgacttcgagtggcttacgcatgcctttg 360
     aatttgcatcttcgcacccgcacagcagccagcaaaagcc 400
                410       420       430       440
     |....|....|....|....|....|....|....|....|
     atcgcaggaaaccctatggatgacccatcaatcaaaaaac 440
     ggttactagatttatatactgacgaacatgaatacgatga 480
     ggttcaagagataccagaagagtcaagcattcagccccg 520
     gagacatcaacgagccataccagcacaaacggatccagcc 560
     actcaggccctggaactgcaactggaccaggagcgacatc 600
                610       620       630       640
     |....|....|....|....|....|....|....|....|
     tgcaggcccgtcagcaggtgcgccacagtcaccagtaatt 640
     gttgtggactcggttcccgagctgccagcgcccaagcaga 680
     aatctgttaagaactcgaagagcaaacagaagcaaaagca 720
     gttggcgaacaaatcgaagattccacgatctccttcgttg 760
     gcgagcagcctgagtagtctggccagcagccttagtggtc 800
                810       820       830       840
     |....|....|....|....|....|....|....|....|
     atagggatcgggacaaagatcgggacaaggatcgggagaa 840
     ccagaacgccgtgccgccgcagacgccgccgttgccaccg 880
     agctacaagcagaaccaaatgaatggtgactctacggctg 920
     cagctggtggtggtgtttctgccccagccacgcccaccac 960
     cgccaataataacaatgcgagccacaacaacggcagcata 1000
                1010      1020      1030      1040
     |....|....|....|....|....|....|....|....|
     atgggcggaggcgtgcaatgaatcaatcggacaactcca 1040
     accccgttctccaggcgccggggggagcgcagtagcctcaa 1080
     tctgactcccctctcccgggatctgagtggtggccacacc 1120
     caggagtccacgacgccagcgaccacgccgagtacgccaa 1160
     gcctagcattaccaaaaaattttcagtatttaaccctgac 1200
                1210      1220      1230      1240
     |....|....|....|....|....|....|....|....|
     tgtgcgaaaagatagcaacggatacgggatgaaggtttcc 1240
     ggagataatcctgtgtttgtggagagcgttaaacccggag 1280
     gcgcagcggagattgcaggcctggttgctggcgatatgat 1320
     actaagggtaaacggccatgaagtgcgactagagaagcat 1360
     ccaactgtagtgggtcttataaaagcctcgacaactgttg 1400
```

FIGURE 17

```
            1410      1420      1430      1440
     |,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|
     agctggcggtgaagcgaagtcagaagctaacgcgaccctc 1440
     ttcggtatcagtggtaacgccctcgacacccattctctct 1480
     ggacgagatcgtaccgcttcaataactgggccacagccgg 1520
     tggacagtattaaacgaagggagatggagacttacaaaat 1560
     tcagacgttgcagaaaatgctggagcaggagaaactaaat 1600
            1610      1620      1630      1640
     |,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|
     ctggagcgattaaaaagcgatcaaaataatccgagctaca 1640
     agctatctgaggcgaatatccgtaagctgcgcgagcaact 1680
     tcatcaagtggggagctgaggatgcaccgactgttaaactt 1720
     caggcggccgctggcaataagaacacagcattactaacgc 1760
     ccaaccaaatccaacacttgtccgcgtccgccactcacag 1800
            1810      1820      1830      1840
     |,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|
     caatcagcaattccatcatcttcaccaccaccacaatctc 1840
     cacaacaacaattatccgccacagcaacagccagctagca 1880
     cctcacccgcattcctgtccctcctgccgcgttcccttttc 1920
     gtccttgtcgctggggcacgcgcaaaaacaagaccgaaaag 1960
     gacctaacgacttcatctccctttggcctaacgacggatt 2000
            2010      2020      2030      2040
     |,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|
     tcctgcagcagcaacggatgagccaccaagcggagtcgat 2040
     gtctcaatcgatgcatcagcacaccagcactccgacctcg 2080
     cagcagttcttccatccgcatcagcaacaacatcgcttta 2120
     aggaaactggaccgacgtcgaaaggcaagaacaagttcct 2160
     aatttcgaggagtttgatcgaggaggatgtgccgccacca 2200
            2210      2220      2230      2240
     |,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|
     ctgccgcagaggaatccgcccaggcagttaaatttggacc 2240
     tgaagaacggaaacgcgtcgccgggtgggtcacatctagt 2280
     ggctccagtttccgatctggatcgcgccacaagtccccaa 2320
     ttaaatagatcccaacaacagcaactgcctaggagcactg 2360
     acaacagcccgagcaatgccaagtccaagcgttccaagat 2400
            2410      2420      2430      2440
     |,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|
     caaaacgaaggccctatcggatcctaagatgtccactcag 2440
     atgctcctgcaaatggaatcggcgagtgcagctggagcag 2480
     ccggaggttccatcgaggtagatggggggtccaccaccgct 2520
     accgcctcgcttgcctggcatgatgacggaggatatgagc 2560
     cgcggcagctgtcagaatctcgctcaacccaattccgtgg 2600
            2610      2620      2630      2640
     |,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|,,,,|
     gcactgccttcaactatcctctggtgtccacaacaacagc 2640
     ggtgcagaacgataacctgaacattgcctttcctttgtcc 2680
     caacgacccaacattgtccagcagctgcagcaatatcaac 2720
     agcagcagcagcatcaaatgagcggtggccaggccacagg 2760
     cgctctgggacagactccgaatttgggaaaaaacaagcat 2800
```

FIGURE 17 cont.

```
          2810        2820        2830        2840
        |___.___|___.___|___.___|___.___|___.___|

        cgacgcgttggttcttcaccggacaatatgcatccacgtc 2840
        atccagatcggataacgaaaacgacttcgggctcgtggga 2880
        gattgttgaaaaggatggcgaatcctccccgcccggaaca 2920
        ccgccgcttccatatctatccagctcccacatgaccgtgc 2960
        tggaagatccgaatgagaacaatcgtggagcagcagcagc 3000

          3010        3020        3030        3040
        |___.___|___.___|___.___|___.___|___.___|

        cggacctggagtctttatcgagtcgcatcagtttacgccg 3040
        atggcgggagcctcttctccgatcccgatatccttacatt 3080
        ccagccacatgcatgcggcccagtcgaacgatacgcagaa 3120
        ggagatcatctcgatggaggacgaaaactcggacttggat 3160
        gagcccttcattgacgagaacggaccctttaacaatctaa 3200

          3210        3220        3230        3240
        |___.___|___.___|___.___|___.___|___.___|

        ctcgtttgttagaggccgagaacgtcactttcctagccat 3240
        cttcctaaactacgtgatctcaaactcggatcccgcgcca 3280
        cttctgttttacctgattactgagttgtacaaggagggca 3320
        cctccaaggacatgcggaaatgggcctacgaaatccactc 3360
        cacattcctcgtgccgcgggctccattgtcatggtatcgc 3400

          3410        3420        3430        3440
        |___.___|___.___|___.___|___.___|___.___|

        caagatgaatcgctggcccgcgaggtggataatgtcctgc 3440
        agttggagtatgacaaagtggagatcctaaggacagtttt 3480
        tctgcgtagccgaaagcgggccaaggacctaatcagtgag 3520
        cagctgcgtgagtttcagcagaagcgcaccgccggcctgg 3560
        gaaccatttacggacccacggacgacaagctggccgaggc 3600

          3610        3620        3630        3640
        |___.___|___.___|___.___|___.___|___.___|

        gaagacggataagctaagggagcaaatcatcgacaaatat 3640
        ttaatgcccaatcttcacgcgcttattgaggatgagaacg 3680
        gttcaccgccggaggatgtacgcaaggtggcgttgtgttc 3720
        tgctctttccaccgtcatctaccgtattttcaacactcgt 3760
        ccgcctccaagcagcatcgttgagcgggtccatcactttg 3800

          3810        3820        3830        3840
        |___.___|___.___|___.___|___.___|___.___|

        tgagcagggacaagagtttcaagtcgcgtattatgggcaa 3840
        aaatcgcaagatgaatgttcgtggtcatccattggtattg 3880
        cgtcaatactatgaagtgacgcactgcaatcattgtcaga 3920
        cgattatctggggcgtgagcccgcaaggttatcattgtac 3960
        agactgtaaattgaacatacaccgtcagtgctcgaaagta 4000

          4010        4020        4030        4040
        |___.___|___.___|___.___|___.___|___.___|

        gtggacgagagttgtcccggtcccttgccccaggcaaaac 4040
        gtctcgcccacaacgacaagatcagtaaattcatgggcaa 4080
        aattcgaccgcgtaccagcgacgtcattggaaatgaaaag 4120
        cgaagtcgtcaagatgaggaattggatgttgagttgactc 4160
        cagaccgtggtcaggcgtcgattgtgcgtcaaccctctga 4200
```

FIGURE 17 cont.

SUBSTITUTE SHEET (RULE 26)

```
                   4210        .4220        4230        4240
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
      .tcggcgaccggatgcgaacatatcgataaggtcgaatgga 4240
      aatacctcctgcaacacttcggggctgaacaccaccgacc  4280
      tgcaaagttcttttcacggcagctgtgccaacgacagtat  4320
      taaccccggcggtggagccggatgcaacatggatttatcc  4360
      acgagcgtggcgtcaacgactccgtcgaccagtggatccg  4400
                   4410        4420        4430        4440
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
      tggcagccggtctgagtgcttttgccgaactgaacgccct  4440
      ggatacagtggataaagaagcgcgtagggagcgttacagt  4480
      cagcatccggagcacaaaagtgcaccagtctctgtgaatc  4520
      ggtcggaatcctacaaggagcgcttgtccaacaagaggaa  4560
      ccgcaacagtcgccgcaagacctctgatccaagcttgtcg  4600
                   4610        4620        4630        4640
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
      tcgcgtcccaatgatgaacaactggacttgggtctatcga  4640
      atgccacttatgtgggcagttcgaattctagtctctcttc  4680
      agctggcggcagcgagagtcccagcacgtcaatggagcac  4720
      tttgctgcacccggagcagcgggtggcgtccaggtgccgc  4760
      caatgggattgaaccagaaccagcaccccatctgctcat   4800
                   4810        4820        4830        4840
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
      ccagcagcacgccccgcagtactgccagcaggattccttt  4840
      caggcgggtttggcaggggccgctgggagcagtgcagcta  4880
      gcaactctagtttctggaatgctggccacccattgcctgt  4920
      agctcgttggacgctggagagcgaggatgaagacgacgtg  4960
      aacgaggcggactggagttccatggtggccgcagaggtgt  5000
                   5010        5020        5030        5040
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
      tggcagccttaacggacgctgagaagaagcgtcaggagat  5040
      tataaatgaaatctatcaaactgaacgcaaccatgtgcgc  5080
      accctaaagctgctggatcgattattcttcctgccactct  5120
      acgagagtggattgctgtcccaggatcatttgctgttgtt  5160
      gttcccgcccgccttgctgtcgctccgtgagattcatggc  5200
                   5210        5220        5230        5240
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
      gccttcgagcagagtcttaagcaacgacgcatcgagcata  5240
      accacgtggtgaacaccatagggatctgctcgccgacat   5280
      gttcgatggtcagtctggagttgttctttgcgagtttgcg  5320
      gctcagttctgtgcccgccagcaaatcgccttggaggcac  5360
      tcaaggagaagcgcaacaaggacgagatgctgcaaaagct  5400
                   5410        5420        5430        5440
      ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
      attgaaaaagtcggagtcacacaaggcatgtcgccgcctc  5440
      gagctaaaggacttgctgcccaccgtgctgcagcgcctca  5480
      ccaagtatccgcttttgtttgagaacctttacaaggtgac  5520
      cgtgcgcttgctgccagaaaacaccaccgaggcggaagcc  5560
      attcaacgggcagtggaatcctctaaaaggattcttgtcg  5600
```

FIGURE 17 cont.

31/34

```
                5610        5620        5630        5640
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        aggtcaaccaggcagtaaagacagcagaggatgctcacaa 5640
        gctgcaaaacattcagcgtaagttagacagatcctcctat 5680
        gacaaggaagagtttaagaaattggacctgacccaacatc 5720
        acctcatccatgacggcaatctgacgatcaagaagaatcc 5760
        tagcgtgcagctacatggacttctgtttgaaaacatgatt 5800

                5810        5820        5830        5840
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        gttttgctgaccaagcaggatgataaatattatctaaaga 5840
        acttgcacaccccgctatcgatcaccaataagccagtcag 5880
        tccaattatgagcattgatgcggcacctttgatccggcag 5920
        gaggcggctgataaaaattcctttttcctcatcaagatga 5960
        agacatcacaaatgttggagctacgcgcgcctagtagctc 6000

                6010        6020        6030        6040
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        ggagtgcaagacatggtttaaacacttctcggatgtggct 6040
        gctcgtcagtctaagaatcgttcaaagaacgcatcaagca 6080
        accatgacacgagtattagtgatccagctctcgccgctat 6120
        tccgcattccaacaccaaagagtcgttggagttgagcact 6160
        gatacagtacagccattggctgcgacagccacattgacta 6200

                6210        6220        6230        6240
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        ccacaccattggccccaatgctgcctatagccacggttac 6240
        accggctccagcgaccaataatagtaacgttagctctctt 6280
        actggagttcagttgcgaaaccctcaacgggatgcgacag 6320
        caagtgaatctgatgcggattatgtaaacacaccaaagcc 6360
        gcgttcgagccaaaatgaagttaatcgcactatgtccata 6400

                6410        6420        6430        6440
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        agaagcactggcgaacccattcagaagtattcggcgaatg 6440
        ggacggaagcaaacgacgttactttacgacactctcagtc 6480
        gactagggaatcggttagaccaggatctactggggaggag 6520
        cgaaactccacgtatggtatggttggaggtaactccaaac 6560
        gcgacagcgccagcattgtctgctcgaacaactcgaacaa 6600

                6610        6620        6630        6640
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        cacgcgcacccttctgatgcagagtcccttggtggaccct 6640
        acggccattcaggtcagcattagtcccgctcacacagcgg 6680
        aacctgtgttgacaccaggagagaagttacgccgcttgga 6720
        cgcctccattaggaatgatttgctggagaagcagaaaatc 6760
        atttgtgatatcttccgtttgccagtggaacactacgacc 6800

                6810        6820        6830        6840
        ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴

        agattgtggacattgccatgatgccagaggcgccgaaaga 6840
        cagtgcagatattgctttagctgcttacgatcagattcaa 6880
        accctgaccaagatgctgaacgagtacatgcacgtcacgc 6920
        ctgagcaagaggtctcagcggtgtccacggcggtttgtgg 6960
        ccactgtcacgagaaggaaaagcttcgtaagaaggtggca 7000
```

FIGURE 17 cont.

SUBSTITUTE SHEET (RULE 26)

```
              7010      7020      7030      7040
         ├┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┤

         ccatcctcatcgttttcttcatcaccaccaccgctgccgc 7040
         ctcccaataggcagcatgcccaggctcaggcccagatacc 7080
         gccatcgcggctgatgcccaaactacaaactcttgatctt 7120
         gacgaagttgccatacacgaagacgatgacggatactgtg 7160
         agatcgacgaactgcgcttaccggctattccgtccaaacc 7200

              7210      7220      7230      7240
         ├┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┤

         acatgagcggcccacaacgccactggctcctttcaatact 7240
         gagccgaaaacttcacaatctgttatagatgcctcgaaac 7280
         gtcaatccactgatgccgttccggagggattactggaaca 7320
         agaaccactcgaaggcgataagacggagaccaagggtgaa 7360
         gataatgaagtgaaaactgtgccgtcagataagctaagtg 7400

              7410      7420      7430      7440
         ├┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┤

         aatcatgcaatgaagagaggcaatgtgtggaggcggatat 7440
         cacaaaggaagtggcagatccaacgacctctaagaatgaa 7480
         gctgcagcatcggtggatgaattaccaagccagagccggg 7520
         agataaaaacggctgaaaacgcaagcaaatctgtagctga 7560
         caaaaaggaagacaacgaggaaaccatcgaagaaggtgtg 7600

              7610      7620      7630      7640
         ├┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┤

         gcatccacggtcgatagctccactcaaacatcaccaactg 7640
         aatcgccgaaagagccggataagttaactggaggatcgag 7680
         cagcacctgtgggccgaatcgcattcagcacgctagtgtg 7720
         ctggagccgagtgtgccctgccatgcactcagcagcattg 7760
         taacaatactgaatgagcagatttccatgcttttgccaaa 7800

              7810      7820      7830      7840
         ├┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┤

         aattaacgaacgcgatatggaaagggagcgattgcgtaaa 7840
         gagaatcaacaccttcgcgagctcttgagtgcgctgcatg 7880
         atcgacagcgagttgatgaagtaaaggaaactccgtttga 7920
         tctaaagaagctgatgcatgctgaggatgtagagtttgac 7960
         gatgatattgacgccatttccaacagttcgctgacgccaa 8000

              8010      8020      8030      8040
         ├┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┤

         cgcctacgccgattcccacagcatcaccaagcgccagcgg 8040
         ccaggtagagacagcggaagccatgaggattactagcact 8080
         gaggatgaggaataggaaacttgtttaatttttttttgta 8120
         taaaatacatatttgttcttgttttttcccgcacgcatttg 8160
         accaattgaattggcggcgtctatgttttacgagtagcct 8200

              8210      8220      8230      8240
         ├┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┼┴┴┴┴┴┴┴┴┴┤

         ttcttagagcatgccttttactactgttaatgtagcttca 8240
         tttgttttcttaaaatattatgtattccacaagagccgat 8280
         aacactttaaattaagtctaacattgtcccctcgtaaaac 8320
         taattaaatacatatatatatgcatattatcgacttaaaa 8360
         aatatcagataaacacacaatgtttgatacatttataaag 8400
```

FIGURE 17 cont.

```
          8410        8420        8430        8440
  ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
aacaataaacaacatagttattgagaagaatagcaaaaaa 8440
aaaataaaaaaaaaaaaaaaaaaa 8464
```

FIGURE 17 cont.

FIGURE 18

PCT

(51) International Patent Classification 6 : C12N 15/12, C07K 14/82, A61K 38/17, G01N 33/68, C12N 15/63, 5/10	A3	(11) International Publication Number: WO 99/47557
		(43) International Publication Date: 23 September 1999 (23.09.99)

(21) International Application Number: PCT/US99/06051

(22) International Filing Date: 18 March 1999 (18.03.99)

(30) Priority Data:
60/078,634   18 March 1998 (18.03.98)   US

(71) Applicant: ONYX PHARMACEUTICALS, INC. [US/US]; 3031 Research Drive, Richmond, CA 94806 (US).

(72) Inventors: BOLLAG, Gideon; 172 Catalina Drive, Hercules, CA 94547 (US). HART, Matthew, J.; 1210 Masonic Avenue, Berkeley, CA 94706 (US). ROSCOE, William; 3099 California Street, San Francisco, CA 94115 (US). POLAKIS, Paul; 509 Barone Lane, Mill Valley, CA 94941 (US). STERNWEIS, Paul; 2103 Flat Creet Drive, Richardson, TX 75080 (US). KOZASA, Tohru; 7415 Centenary, Dallas, TX 75225 (US). JIANG, Xuejun; 2400 Waterview Parkway #932, Richardson, TX 75080 (US).

(74) Agent: GIOTTA, Gregory; Onyx Pharmaceuticals, Inc., 3031 Research Drive, Richmond, CA 94806 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
18 November 1999 (18.11.99)

---

(54) Title: IDENTIFICATION OF FACTORS WHICH MEDIATE THE INTERACTION OF HETEROTRIMERIC G PROTEINS AND MONOMERIC G PROTEINS

(57) Abstract

Monomeric GTPase guanine nucleotide exchange factor (GEF) have been identified which also contain an RGS region analogous to those of GTPase activating proteins (GAP). One of these GEF proteins, a Rho GEF has been demonstrated to contain an RGS sequence that has GAP activity toward a α subunit of a heterotrimeric G protein.

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6   C12N15/12   C07K14/82   A61K38/17   G01N33/68   C12N15/63
C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6   C12N   C07K   A61K   G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MATTHEW J. HART ET AL.: "Identification of a novel Guanine nucleotide exchange factor for the Rho GTPase" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 271, no. 41, 11 October 1996 (1996-10-11), pages 25452-25458, XP002057776 MD US cited in the application the whole document —/—	1-28

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 September 1999	05/10/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl. Fax: (+31–70) 340–3016	Montero Lopez, B

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	DAVID M. BERMAN ET AL.: "GAIP and RGS4 are GTPase-activating proteins for the Gi subfamily of G protein alpha subunits" CELL, vol. 86, no. 3, 9 August 1996 (1996-08-09), pages 445-452, XP002111126 NA US cited in the application the whole document	1-28
P,X	KOZASA T ET AL: "p115 RhoGEF, a GTPase activating protein for Galpha12 and Galpha13 'see comments!." SCIENCE, (1998 JUN 26) 280 (5372) 2109-11. , XP002111127 abstract page 2109, right-hand column, paragraph 2 -page 2111, middle column, paragraph 2	1-28
P,X	ALAN HALL: "G proteins and small GTPases: Distant relatives keep in touch" SCIENCE, vol. 280, 26 June 1998 (1998-06-26), pages 2074-2075, XP002111128 DC page 2074, middle column, paragraph 2 -page 2075, right-hand column, paragraph 2	1-28
P,X	HART M J ET AL: "Direct stimulation of the guanine nucleotide exchange activity of p115 RhoGEF by Galpha13 'see comments!." SCIENCE, (1998 JUN 26) 280 (5372) 2112-4. , XP002111129 the whole document	1-28
P,X	MAO J ET AL: "Guanine nucleotide exchange factor GEF115 specifically mediates activation of Rho and serum response factor by the G protein alpha subunit Galpha13." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1998 OCT 27) 95 (22) 12973-6. , XP002111130 abstract page 12974, right-hand column, paragraph 2 -page 12976, right-hand column, paragraph 1	1-29

-/--

3

	Intern~~ational~~ Application No
	PCT/uS 99/06051

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 98 20127 A (ONYX PHARMACEUTICALS) 14 May 1998 (1998-05-14) page 2, line 13 -page 9, line 32 page 19, line 13 -page 23, line 2 page 36, line 8 -page 37, line 32 ---	19-25
P,X	WO 98 57990 A (ONYX PHARMACEUTICALS) 23 December 1998 (1998-12-23) page 3, line 16 -page 4, line 2 page 4, line 19 -page 5, line 13 page 6, line 20 -page 7, line 3 page 8, line 26 -page 9, line 12 page 19, line 14 -page 20, line 11 page 20, line 24 -page 24, line 30 ---	19-25
T	FUKUHARA S ET AL: "A novel PDZ domain containing guanine nucleotide exchange factor links heterotrimeric G proteins to Rho." JOURNAL OF BIOLOGICAL CHEMISTRY, (1999 FEB 26) 274 (9) 5868-79. , XP002111131 abstract page 5873, left-hand column, last paragraph -page 5874, right-hand column, paragraph 3 page 5875, right-hand column, paragraph 2 -page 5878, right-hand column, last paragraph -----	

3

INTERNATIONAL SEARCH REPORT

Box I   Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
   because they relate to subject matter not required to be searched by this Authority, namely:
   > Remark: Although claim 18
   > is directed to a method of treatment of the human/animal
   > body, the search has been carried out and based on the alleged
   > effects of the compound/composition.

2. [ ] Claims Nos.:
   because they relate to parts of the International Application that do not comply with the prescribed requirements to such
   an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II   Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all
   searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment
   of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report
   covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is
   restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (1)) (July 1998)

INTERNATIONAL SEARCH REPORT

.formation on patent family members

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
WO 9820127	A	14-05-1998	AU	4801097 A		29-05-1998
WO 9857990	A	23-12-1998	AU	7966498 A		04-01-1999